PTH MOMENT NOISE-TO-STATE STABILITY OF STOCHASTIC DIFFERENTIAL EQUATIONS WITH PERSISTENT NOISE*

DAVID MATEOS-NÚÑEZ AND JORGE CORTÉS†

Abstract. This paper studies the stability properties of stochastic differential equations subject to persistent noise (including the case of additive noise), which is noise that is present even at the equilibria of the underlying differential equation and does not decay with time. The class of systems we consider exhibit disturbance attenuation outside a closed, not necessarily bounded, set. We identify conditions, based on the existence of Lyapunov functions, to establish the noise-to-state stability in probability and in pth moment of the system with respect to a closed set. As part of our analysis, we study the concept of two functions being proper with respect to each other formalized via pair of inequalities with comparison functions. We show that such inequalities define several equivalence relations for increasingly strong refinements on the comparison functions. We also provide a complete characterization of the properties that a pair of functions must satisfy to belong to the same equivalence class. This characterization allows us to provide checkable conditions to determine whether a function satisfies the requirements to be a strong NSS-Lyapunov function in probability or a pth moment NSS-Lyapunov function. Several examples illustrate our results.

1. Introduction.

Stochastic differential equations (SDEs) go beyond ordinary differential equations (ODEs) to deal with systems subject to stochastic perturbations of a particular type, known as white noise. Applications are numerous and include option pricing in the stock market, networked systems with noisy communication channels, and, in general, scenarios whose complexity cannot be captured by deterministic models. In this paper, we study SDEs subject to *persistent noise* (including the case of additive noise), i.e., systems for which the noise is present even at the equilibria of the underlying ODE and does not decay with time. Such scenarios arise, for instance, in control-affine systems when the input is corrupted by persistent noise. For these systems, the presence of persistent noise makes it impossible to establish in general a stochastic notion of asymptotic stability for the (possibly unbounded) set of equilibria of the underlying ODE. Our aim here is to develop notions and tools to study the stability properties of these systems and provide probabilistic guarantees on the size of the state of the system.

Literature review: In general, it is not possible to obtain explicit descriptions of the solutions of SDEs. Fortunately, the Lyapunov techniques used to study the qualitative behavior of ODEs [6, 10] can be adapted to study the stability properties of SDEs as well [7, 13, 27]. Depending on the notion of stochastic convergence used, there are several types of stability results in SDEs. These include stochastic stability (or stability in probability), stochastic asymptotic stability, almost sure exponential stability, and pth moment asymptotic stability, see e.g., [13, 14, 26, 27]. However, these notions are not appropriate in the presence of persistent noise because they require the effect of the noise on the set of equilibria to either vanish or decay with time. To deal with persistent noise, as well as other system properties like delays, a concept of ultimate boundedness is required that generalizes the notion of convergence. As an

*A preliminary version of this manuscript was presented as [15] at the 2013 American Control Conference, Washington, D.C. This manuscript is a revision of the version submitted to SIAM Journal on Control and Optimization 52 (4) (2014), 2399-2421.

†The authors are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA, {dmateosn,cortes}@ucsd.edu.

example, for stochastic delay differential equations, [28] considers a notion of ultimate bound in pth moment [21] and employs Lyapunov techniques to establish it. More generally, for mean-square random dynamical systems, the concept of forward attractor [9] describes a notion of convergence to a dynamic neighborhood and employs contraction analysis to establish it. Similar notions of ultimate boundedness for the state of a system, now in terms of the size of the disturbance, are also used for differential equations, and many of these notions are inspired by dissipativity properties of the system that are captured via *dissipation inequalities* of a suitable Lyapunov function: such inequalities encode the fact that the Lyapunov function decreases along the trajectories of the system as long as the state is "big enough" with regards to the disturbance. As an example, the concept of input-to-state stability (ISS) goes hand in hand with the concept of ISS-Lyapunov function, since the existence of the second implies the former (and, in many cases, a converse result is also true [24]). Along these lines, the notion of practical stochastic input-to-state stability (SISS) in [12, 29] generalizes the concept of ISS to SDEs where the disturbance or input affects both the deterministic term of the dynamics and the diffusion term modeling the role of the noise. Under this notion, the state bound is guaranteed in probability, and also depends, as in the case of ISS, on a decaying effect of the initial condition plus an increasing function of the sum of the size of the input and a positive constant related to the persistent noise. For systems where the input modulates the covariance of the noise, SISS corresponds to noise-to-state-stability (NSS) [3], which guarantees, in probability, an ultimate bound for the state that depends on the magnitude of the noise covariance. That is, the noise in this case plays the main role, since the covariance can be modulated arbitrarily and can be unknown. This is the appropriate notion of stability for the class of SDEs with persistent noise considered in this paper, which are nonlinear systems affine in the input, where the input corresponds to white noise with locally bounded covariance. Such systems cannot be studied under the ISS umbrella, because the stochastic integral against Brownian motion has infinite variation, whereas the integral of a legitimate input for ISS must have finite variation.

Statement of contributions: The contributions of this paper are twofold. Our first contribution concerns the noise-to-state stability of systems described by SDEs with persistent noise. We generalize the notion of noise-dissipative Lyapunov function, which is a positive semidefinite function that satisfies a dissipation inequality that can be nonexponential (by this we mean that the inequality admits a convex $\mathcal{K}_\infty$ gain instead of the linear gain characteristic of exponential dissipativity). We also introduce the concept of $p$thNSS-Lyapunov function with respect to a closed set, which is a noise-dissipative Lyapunov function that in addition is proper with respect to the set with a convex lower-bound gain function. Using this framework, we show that noise-dissipative Lyapunov functions have NSS dynamics and we characterize the overshoot gain. More importantly, we show that the existence of a $p$thNSS-Lyapunov function with respect to a closed set implies that the system is NSS in pth moment with respect to the set. Our second contribution is driven by the aim of providing alternative, structured ways to check the hypotheses of the above results. We introduce the notion of two functions being proper with respect to each other as a generalization of the notion of properness with respect to a set. We then develop a methodology to verify whether two functions are proper with respect to each other by analyzing the associated pair of inequalities with increasingly strong refinements that involve the classes $\mathcal{K}$, $\mathcal{K}_\infty$, and $\mathcal{K}_\infty$ plus a convexity property. We show that these refinements define equivalence relations between pairs of functions, thereby pro-

ducing nested partitions on the space of functions. This provides a useful way to deal with these inequalities because the construction of the gains is explicit when the transitivity property is exploited. This formalism motivates our characterization of positive semidefinite functions that are proper, in the various refinements, with respect to the Euclidean distance to their nullset. This characterization is technically challenging because we allow the set to be noncompact, and thus the pre-comparison functions can be discontinuous. We devote special attention to the case when the set is a subspace and examine the connection with seminorms. Finally, we show how this framework allows us to develop an alternative formulation of our stability results.

Organization: The paper is organized as follows. Section 2 introduces preliminaries on seminorms, comparison functions, and SDEs. Section 3 presents the NSS stability results and Section 4 develops the methodology to help verify their hypotheses. Finally, Section 5 discusses our conclusions and ideas for future work.

2. Preliminary notions.

This section reviews some notions on comparison functions and stochastic differential equations that are used throughout the paper.

2.1. Notational conventions.

Let $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ be the sets of real and nonnegative real numbers, respectively. We denote by $\mathbb{R}^n$ the n-dimensional Euclidean space. A subspace $\mathcal{U} \subseteq \mathbb{R}^n$ is a subset which is also a vector space. Given a matrix $A \in \mathbb{R}^{m\times n}$, its nullspace $\mathcal{N}(A) \triangleq \{x \in \mathbb{R}^n : Ax = 0\}$ is a subspace. Given $\mathcal{D} \subseteq \mathbb{R}^n$, we denote by $\mathcal{C}(\mathcal{D};\mathbb{R}_{\geq 0})$ and $\mathcal{C}^2(\mathcal{D};\mathbb{R}_{\geq 0})$ the set of positive semidefinite functions defined on $\mathcal{D}$ that are continuous and continuously twice differentiable (if $\mathcal{D}$ is open), respectively. Given $\mathrm{V} \in \mathcal{C}^2(\mathbb{R}^n;\mathbb{R}_{\geq 0})$, we denote its gradient by $\nabla \mathrm{V}$ and its Hessian by $\nabla^2\mathrm{V}$. A seminorm is a function $S : \mathbb{R}^n \to \mathbb{R}$ that is positively homogeneous, i.e., $S(\lambda x) = |\lambda|S(x)$ for any $\lambda \in \mathbb{R}$, and satisfies the triangular inequality, i.e., $S(x + y) \leq S(x) + S(y)$ for any $x, y \in \mathbb{R}^n$. From these properties it can be deduced that $S \in \mathcal{C}(\mathbb{R}^n;\mathbb{R}_{\geq 0})$ and its nullset is always a subspace. If, moreover, the function S is positive definite, i.e., $S(x) = 0$ implies $x = 0$, then S is a norm. The Euclidean norm of $x \in \mathbb{R}^n$ is denoted by $\|x\|_2$, and the Frobenius norm of the matrix $A \in \mathbb{R}^{m\times n}$ is $|A|_{\mathcal{F}} \triangleq \sqrt{\mathrm{trace}\,(A^T A)} = \sqrt{\mathrm{trace}\,(AA^T)}$. For any matrix $A \in \mathbb{R}^{m\times n}$, the function $\|x\|_A \triangleq \|Ax\|_2$ is a seminorm and can be viewed as a distance to $\mathcal{N}(A)$. For a symmetric positive semidefinite real matrix $A \in \mathbb{R}^{n\times n}$, we order its eigenvalues as $\lambda_{\max}(A) \triangleq \lambda_1(A) \geq \cdots \geq \lambda_n(A) \triangleq \lambda_{\min}(A)$, so if the dimension of $\mathcal{N}(A)$ verifies $\dim(\mathcal{N}(A)) = k \leq n$, then $\lambda_{n-k}(A)$ is the minimum nonzero eigenvalue of A. The Euclidean distance from x to a set $\mathcal{U} \subseteq \mathbb{R}^n$ is defined by $|x|_{\mathcal{U}} \triangleq \inf\{\|x-u\|_2 : u \in \mathcal{U}\}$. The function $|.|_{\mathcal{U}}$ is continuous when $\mathcal{U}$ is closed. Given $f, g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, we say that $f(s)$ is in $\mathcal{O}(g(s))$ as $s \to \infty$ if there exist constants $\kappa, s_0 > 0$ such that $f(s) < \kappa g(s)$ for all $s > s_0$.

2.2. Comparison, convex, and concave functions.

Here we introduce some classes of comparison functions following [6] that are useful in our technical treatment. A continuous function $\alpha : [0, b) \to \mathbb{R}_{\geq 0}$, for $b > 0$ or $b = \infty$, is class $\mathcal{K}$ if $\alpha(0) = 0$ and is strictly increasing. A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is class $\mathcal{K}_\infty$ if $\alpha \in \mathcal{K}$ and is unbounded. A continuous function $\mu : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is class $\mathcal{KL}$ if, for each fixed $s \geq 0$, the function $r \mapsto \mu(r, s)$ is class $\mathcal{K}$, and, for each fixed $r \geq 0$, the function $s \mapsto \mu(r, s)$ is decreasing and $\lim_{s\to\infty} \mu(r, s) = 0$. If α_1, α_2 are class $\mathcal{K}$ and the domain of α_1 contains the range of α_2, then their composition $\alpha_1 \circ \alpha_2$ is class $\mathcal{K}$ too. If α_3, α_4 are class $\mathcal{K}_\infty$, then both the inverse function α_3^{-1} and their composition $\alpha_3 \circ \alpha_4$ are

class $\mathcal{K}_\infty$. In our technical treatment, it is sometimes convenient to require comparison functions to satisfy additional convexity properties. A real-valued function f defined in a convex set X in a vector space is convex if $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$ for each $x, y \in X$ and any $\lambda \in [0,1]$, and is concave if $-f$ is convex. By [2, Ex. 3.3], if $f : [a,b] \to [f(a), f(b)]$ is a strictly increasing convex (respectively, concave) function, then the inverse function $f^{-1} : [f(a), f(b)] \to [a,b]$ is strictly increasing and concave (respectively, convex). Also, following [2, Section 3], if $f, g : \mathbb{R} \to \mathbb{R}$ are convex (respectively, concave) and f is nondecreasing, then the composition $f \circ g$ is also convex (respectively, concave).

2.3. Brownian motion. We review some basic facts on probability and introduce the notion of Brownian motion following [14]. Throughout the paper, we assume that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ is a complete probability space, where $\mathbb{P}$ is a probability measure defined on the σ-algebra $\mathcal{F}$, which contains all the subsets of Ω of probability 0. The filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is a family of sub-σ-algebras of $\mathcal{F}$ satisfying $\mathcal{F}_t \subseteq \mathcal{F}_s \subseteq \mathcal{F}$ for any $0 \leq t < s < \infty$; we assume it is right continuous, i.e., $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$ for any $t \geq 0$, and $\mathcal{F}_0$ contains all the subsets of Ω of probability 0. The Borel σ-algebra in $\mathbb{R}^n$, denoted by $\mathcal{B}^n$, or in $[t_0, \infty)$, denoted by $\mathcal{B}([t_0,\infty))$, are the smallest σ-algebras that contain all the open sets in $\mathbb{R}^n$ or $[t_0, \infty)$, respectively. A function $X : \Omega \to \mathbb{R}^n$ is $\mathcal{F}$-measurable if the set $\{\omega \in \Omega : X(\omega) \in A\}$ belongs to $\mathcal{F}$ for any $A \in \mathcal{B}^n$. We call such function a ($\mathcal{F}$-measurable) $\mathbb{R}^n$-valued random variable. If X is a real-valued random variable that is integrable with respect to $\mathbb{P}$, its expectation is $\mathbb{E}[X] = \int_\Omega X(\omega)\mathrm{d}\mathbb{P}(\omega)$. A function $f : \Omega \times [t_0, \infty) \to \mathbb{R}^n$ is $\mathcal{F} \times \mathcal{B}$-measurable (or just measurable) if the set $\{(\omega, t) \in \Omega \times [t_0,\infty) : f(\omega, t) \in A\}$ belongs to $\mathcal{F} \times \mathcal{B}([t_0, \infty))$ for any $A \in \mathcal{B}^n$. We call such function an $\{\mathcal{F}_t\}$-adapted process if $f(., t) : \Omega \to \mathbb{R}^n$ is $\mathcal{F}_t$-measurable for every $t \geq t_0$. At times, we omit the dependence on "ω", in the sense that we refer to the indexed family of random variables, and refer to the random process $f = \{f(t)\}_{t\geq t_0}$. We define $\mathcal{L}^1([t_0, \infty); \mathbb{R}^n)$ as the set of all $\mathbb{R}^n$-valued measurable $\{\mathcal{F}_t\}$-adapted processes f such that $\mathbb{P}(\{\omega \in \Omega : \int_{t_0}^T \|f(\omega, s)\|_2 \, \mathrm{d}s < \infty\}) = 1$ for every $T > t_0$. Similarly, $\mathcal{L}^2([t_0, \infty); \mathbb{R}^{n\times m})$ denotes the set of all $\mathbb{R}^{n\times m}$-matrix-valued measurable $\{\mathcal{F}_t\}$-adapted processes G such that $\mathbb{P}(\{\omega \in \Omega : \int_{t_0}^T |G(\omega, s)|_{\mathcal{F}}^2 \, \mathrm{d}s < \infty\}) = 1$ for every $T > t_0$.

A one-dimensional Brownian motion $\mathrm{B} : \Omega \times [t_0, \infty) \to \mathbb{R}$ defined in the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ is an $\{\mathcal{F}_t\}$-adapted process such that

- $\mathbb{P}(\{\omega \in \Omega : \mathrm{B}(\omega, t_0) = 0\}) = 1$;
- the mapping $\mathrm{B}(\omega, .) : [t_0, \infty) \to \mathbb{R}$, called sample path, is continuous also with probability 1;
- the increment $\mathrm{B}(., t) - \mathrm{B}(., s) : \Omega \to \mathbb{R}$ is independent of $\mathcal{F}_s$ for $t_0 \leq s < t < \infty$ (i.e., if $S_b \triangleq \{\omega \in \Omega : \mathrm{B}(\omega, t) - \mathrm{B}(\omega, s) \in (-\infty, b)\}$, for $b \in \mathbb{R}$, then $\mathbb{P}(A \cap S_b) = \mathbb{P}(A)\mathbb{P}(S_b)$ for all $A \in \mathcal{F}_s$ and all $b \in \mathbb{R}$). In addition, this increment is normally distributed with zero mean and variance $t - s$.

An m-dimensional Brownian motion $\mathrm{B} : \Omega \times [t_0, \infty) \to \mathbb{R}^m$ is given by $\mathrm{B}(\omega, t) = [\mathrm{B}_1(\omega, t), \ldots, \mathrm{B}_m(\omega, t)]^T$, where each B_i is a one-dimensional Brownian motion and, for each $t \geq t_0$, the random variables $\mathrm{B}_1(t), ..., \mathrm{B}_m(t)$ are independent.

2.4. Stochastic differential equations. Here we review some basic notions on stochastic differential equations (SDEs) following [14]; other useful references are [7,

17, 18]. Consider the n-dimensional SDE

$$\mathrm{d}x(\omega,t) = f\big(x(\omega,t),t\big)\mathrm{d}t + G\big(x(\omega,t),t\big)\Sigma(t)\,\mathrm{dB}(\omega,t), \tag{2.1}$$

where $x(\omega,t) \in \mathbb{R}^n$ is a realization at time t of the random variable $x(.,t) : \Omega \to \mathbb{R}^n$, for $t \in [t_0, \infty)$. The initial condition is given by $x(\omega, t_0) = x_0$ with probability 1 for some $x_0 \in \mathbb{R}^n$. The functions $f : \mathbb{R}^n \times [t_0,\infty) \to \mathbb{R}^n$, $G : \mathbb{R}^n \times [t_0,\infty) \to \mathbb{R}^{n\times q}$, and $\Sigma : [t_0,\infty) \to \mathbb{R}^{q\times m}$ are measurable. The functions f and G are regarded as a model for the architecture of the system and, instead, Σ is part of the model for the stochastic disturbance; at any given time Σ determines a linear transformation of the m-dimensional Brownian motion $\{\mathrm{B}(t)\}_{t\geq t_0}$, so that at time $t \geq t_0$ the input to the system is the process $\{\Sigma(t)\mathrm{B}(t)\}_{t\geq t_0}$, with covariance $\int_{t_0}^t \Sigma(t)\Sigma(t)^T \mathrm{d}s$. The distinction between the roles of G and Σ is irrelevant for the SDE; both together determine the effect of the Brownian motion. The integral form of (2.1) is given by

$$x(\omega,t) = x_0 + \int_{t_0}^t f\big(x(\omega,s),s\big)\mathrm{d}s + \int_{t_0}^t G\big(x(\omega,s),s\big)\Sigma(s)\,\mathrm{dB}(\omega,s),$$

where the second integral is an stochastic integral [14, p. 18]. A $\mathbb{R}^n$-valued random process $\{x(t)\}_{t\geq t_0}$ is a solution of (2.1) with initial value x_0 if

(i) is continuous with probability 1, $\{\mathcal{F}_t\}$-adapted, and satisfies $x(\omega,t_0) = x_0$ with probability 1,
(ii) the processes $\{f(x(t),t)\}_{t\geq t_0}$ and $\{G(x(t),t)\}_{t\geq t_0}$ belong to $\mathcal{L}^1([t_0,\infty);\mathbb{R}^n)$ and $\mathcal{L}^2([t_0,\infty);\mathbb{R}^{n\times m})$ respectively, and
(iii) equation (2.1) holds for every $t \geq t_0$ with probability 1.

A solution $\{x(t)\}_{t\geq t_0}$ of (2.1) is unique if any other solution $\{\bar{x}(t)\}_{t\geq t_0}$ with $\bar{x}(t_0) = x_0$ differs from it only in a set of probability 0, that is, $\mathbb{P}(\big\{x(t) = \bar{x}(t) \;\; \forall\, t \geq t_0\big\}) = 1$.

We make the following assumptions on the objects defining (2.1) to guarantee existence and uniqueness of solutions.

Assumption 2.1. *We assume Σ is essentially locally bounded. Furthermore, for any $T > t_0$ and $n \geq 1$, we assume there exists $K_{T,n} > 0$ such that, for almost every $t \in [t_0, T]$ and all $x, y \in \mathbb{R}^n$ with* $\max\big\{\|x\|_2, \|y\|_2\big\} \leq n$,

$$\max\big\{\, \|f(x,t) - f(y,t)\|_2^2\,,\, |G(x,t) - G(y,t)|_{\mathcal{F}}^2\,\big\} \leq K_{T,n}\|x - y\|_2^2.$$

Finally, we assume that for any $T > t_0$, there exists $K_T > 0$ such that, for almost every $t \in [t_0,T]$ and all $x \in \mathbb{R}^n$, $x^T f(x,t) + \frac{1}{2}|G(x,t)|_{\mathcal{F}}^2 \leq K_T(1 + \|x\|_2^2)$.

According to [14, Th. 3.6, p. 58], Assumption 2.1 is sufficient to guarantee global existence and uniqueness of solutions of (2.1) for each initial condition $x_0 \in \mathbb{R}^n$.

We conclude this section by presenting a useful operator in the stability analysis of SDEs. Given a function $\mathrm{V} \in \mathcal{C}^2(\mathbb{R}^n;\mathbb{R}_{\geq 0})$, we define the generator of (2.1) acting on the function V as the mapping $\mathcal{L}[\mathrm{V}] : \mathbb{R}^n \times [t_0,\infty) \to \mathbb{R}$ given by

$$\mathcal{L}[\mathrm{V}](x,t) \triangleq \nabla\mathrm{V}(x)^T f(x,t) + \tfrac{1}{2}\,\mathrm{trace}\,\Big(\Sigma(t)^T G(x,t)^T \nabla^2\mathrm{V}(x) G(x,t)\Sigma(t)\Big). \tag{2.2}$$

It can be shown that $\mathcal{L}[\mathrm{V}](x,t)$ gives the expected rate of change of V along a solution of (2.1) that passes through the point x at time t, so it is a generalization of the Lie

derivative. According to [14, Th. 6.4, p. 36], if we evaluate V along the solution $\{x(t)\}_{t\geq t_0}$ of (2.1), then the process $\{\mathrm{V}(x(t))\}_{t\geq t_0}$ satisfies the new SDE

$$\mathrm{V}(x(t)) = \mathrm{V}(x_0) + \int_{t_0}^{t} \mathcal{L}[\mathrm{V}](x(s),s)\mathrm{d}s + \int_{t_0}^{t} \nabla \mathrm{V}(x(s))^T G(x(s),s)\Sigma(s)\mathrm{dB}(s). \tag{2.3}$$

Equation (2.3) is known as Itô's formula and corresponds to the stochastic version of the chain rule.

3. Noise-to-state stability via noise-dissipative Lyapunov functions.

In this section, we study the stability of stochastic differential equations subject to persistent noise. Our first step is the introduction of a novel notion of stability. This captures the behavior of the pth moment of the distance (of the state) to a given closed set, as a function of two objects: the initial condition and the maximum size of the covariance. After this, our next step is to derive several Lyapunov-type stability results that help determine whether a stochastic differential equation enjoys these stability properties. The following definition generalizes the concept of noise-to-state stability given in [3].

Definition 3.1. (Noise-to-state stability with respect to a set). *The system* (2.1) *is* noise-to-state stable (NSS) in probability with respect to the set $\mathcal{U} \subseteq \mathbb{R}^n$ *if for any* $\epsilon > 0$ *there exist* $\mu \in \mathcal{KL}$ *and* $\theta \in \mathcal{K}$ *(that might depend on* ϵ*), such that*

$$\mathbb{P}\Big\{ |x(t)|_{\mathcal{U}}^p > \mu\big(|x_0|_{\mathcal{U}}, t-t_0\big) + \theta\Big(\operatorname*{ess\,sup}_{t_0\leq s\leq t} |\Sigma(s)|_{\mathcal{F}} \Big) \Big\} \leq \epsilon, \tag{3.1}$$

for all $t \geq t_0$ *and any* $x_0 \in \mathbb{R}^n$*. And the system* (2.1) *is* pth moment noise-to-state stable ($p$thNSS) with respect to $\mathcal{U}$ *if there exist* $\mu \in \mathcal{KL}$ *and* $\theta \in \mathcal{K}$*, such that*

$$\mathbb{E}\big[|x(t)|_{\mathcal{U}}^p\big] \leq \mu\big(|x_0|_{\mathcal{U}}, t-t_0\big) + \theta\Big(\operatorname*{ess\,sup}_{t_0\leq s\leq t} |\Sigma(s)|_{\mathcal{F}} \Big), \tag{3.2}$$

for all $t \geq t_0$ *and any* $x_0 \in \mathbb{R}^n$*. The gain functions* μ *and* θ *are the* overshoot gain *and the* noise gain*, respectively.*

The quantity $|\Sigma(t)|_{\mathcal{F}} = \sqrt{\mathrm{trace}\left(\Sigma(t)\Sigma(t)^T\right)}$ is a measure of the size of the noise because it is related to the infinitesimal covariance $\Sigma(t)\Sigma(t)^T$. The choice of the pth power is irrelevant in the statement in probability since one could take any $\mathcal{K}_\infty$ function evaluated at $|x(t)|_{\mathcal{U}}$. However, this would make a difference in the statement in expectation. (Also, we use the same power for convenience.) When the set $\mathcal{U}$ is a subspace, we can substitute $|.|_{\mathcal{U}}$ by $\|.\|_A$, for some matrix $A \in \mathbb{R}^{m\times n}$ with $\mathcal{N}(A) = \mathcal{U}$. In such a case, the definition above does not depend on the choice of the matrix A.

Remark 3.2. (NSS is not a particular case of ISS). The concept of NSS is not a particular case of input-to-state stability (ISS) [23] for systems that are affine in the input, namely,

$$\dot{y} = f(y,t) + G(y,t)u(t) \;\;\Leftrightarrow\;\; y(t) = y(t_0) + \int_{t_0}^{t} f(y(s),s)\,\mathrm{d}s + \int_{t_0}^{t} G(y(s),s)u(s)\,\mathrm{d}s,$$

where $u : [t_0,\infty) \to \mathbb{R}^q$ is measurable and essentially locally bounded [22, Sec. C.2]. The reason is the following: the components of the vector-valued function

$\int_{t_0}^t G(y(s),s)u(s)\,\mathrm{d}s$ are differentiable almost everywhere by the Lebesgue fundamental theorem of calculus [16, p. 289], and thus absolutely continuous [16, p. 292] and with bounded variation [16, Prop. 8.5]. On the other hand, at any time previous to $t_k(t) \triangleq \min\{t, \inf\{s \geq t_0 : \|x(s)\|_2 \geq k\}\}$, the driving disturbance of (2.1) is the vector-valued function $\int_{t_0}^{t_k(t)} G(x(s),s)\Sigma(s)\mathrm{dB}(s)$, whose ith component has quadratic variation [14, Th. 5.14, p. 25] equal to

$$\int_{t_0}^{t_k(t)} \sum_{j=1}^{m} |\sum_{l=1}^{q} G(x(s),s)_{il}\Sigma(s)_{lj}|^2 \mathrm{d}s > 0.$$

Since a continuous process that has positive quadratic variation must have infinite variation [8, Th. 1.10], we conclude that the driving disturbance in this case is not allowed in the ISS framework. •

Our first goal now is to provide tools to establish whether a stochastic differential equation enjoys the noise-to-state stability properties given in Definition 3.1. To achieve this, we look at the dissipativity properties of a special kind of energy functions along the solutions of (2.1).

Definition 3.3. (Noise-dissipative Lyapunov function). *A function* $\mathrm{V} \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R}_{\geq 0})$ *is a* noise-dissipative Lyapunov function *for* (2.1) *if there exist* $\mathrm{W} \in \mathcal{C}(\mathbb{R}^n; \mathbb{R}_{\geq 0})$, $\sigma \in \mathcal{K}$, *and concave* $\eta \in \mathcal{K}_\infty$ *such that*

$$\mathrm{V}(x) \leq \eta(\mathrm{W}(x)), \tag{3.3}$$

for all $x \in \mathbb{R}^n$, *and the following dissipation inequality holds:*

$$\mathcal{L}[\mathrm{V}](x,t) \leq -\mathrm{W}(x) + \sigma\big(|\Sigma(t)|_{\mathcal{F}}\big), \tag{3.4}$$

for all $(x,t) \in \mathbb{R}^n \times [t_0, \infty)$.

Remark 3.4. (Itô formula and exponential dissipativity). Interestingly, the conditions (3.3) and (3.4) are equivalent to

$$\mathcal{L}[\mathrm{V}](x,t) \leq -\eta^{-1}(\mathrm{V}(x)) + \sigma\big(|\Sigma(t)|_{\mathcal{F}}\big), \tag{3.5}$$

for all $x \in \mathbb{R}^n$, where $\eta^{-1} \in \mathcal{K}_\infty$ is convex. Note that, since $\mathcal{L}[\mathrm{V}]$ is not the Lie derivative of V (as it contains the Hessian of V), one cannot directly deduce from (3.5) the existence of a continuously twice differentiable function $\tilde{\mathrm{V}}$ such that

$$\mathcal{L}[\tilde{\mathrm{V}}](x,t) \leq -c\tilde{\mathrm{V}}(x) + \tilde{\sigma}\big(|\Sigma(t)|_{\mathcal{F}}\big), \tag{3.6}$$

as instead can be done in the context of ISS, see e.g. [19]. •

Example 3.5. (A noise-dissipative Lyapunov function). Assume that $h : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and verifies

$$\gamma(\|x - x'\|_2^2) \leq (x - x')^T(\nabla h(x) - \nabla h(x')) \tag{3.7}$$

for some convex function $\gamma \in \mathcal{K}_\infty$ for all $x, x' \in \mathbb{R}^n$. In particular, this implies that h is strictly convex. (Incidentally, any strongly convex function verifies (3.7) for some choice of γ linear and strictly increasing.) Consider now the dynamics

$$\mathrm{d}x(\omega,t) = -\big(\delta \mathsf{L} x(\omega,t) + \nabla h(x(\omega,t))\big)\mathrm{d}t + \Sigma(t)\,\mathrm{dB}(\omega,t), \tag{3.8}$$

for all $t \in [t_0, \infty)$, where $x(\omega, t_0) = x_0$ with probability 1 for some $x_0 \in \mathbb{R}^n$, and $\delta > 0$. Here, the matrix $\mathsf{L} \in \mathbb{R}^{n\times n}$ is symmetric and positive semidefinite, and the matrix-valued function $\Sigma : [t_0, \infty) \to \mathbb{R}^{n\times m}$ is continuous. This dynamics corresponds to the SDE (2.1) with $f(x,t) \triangleq -\delta \mathsf{L} x - \nabla h(x)$ and $G(x,t) \triangleq \mathrm{I}_n$ for all $(x,t) \in \mathbb{R}^n \times [t_0, \infty)$.

Let $x^* \in \mathbb{R}^n$ be the unique solution of the Karush-Kuhn-Tucker [2] condition $\delta \mathsf{L} x^* = -\nabla h(x^*)$, corresponding to the unconstrained minimization of $F(x) \triangleq \frac{\delta}{2} x^T \mathsf{L} x + h(x)$. Consider then the candidate Lyapunov function $\mathrm{V} \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R}_{\geq 0})$ given by $\mathrm{V}(x) \triangleq \frac{1}{2}(x - x^*)^T(x - x^*)$. Using (2.2), we obtain that, for all $x \in \mathbb{R}^n$,

$$
\begin{aligned}
\mathcal{L}[\mathrm{V}](x,t) &= -(x - x^*)^T \Big(\delta \mathsf{L} x + \nabla h(x)\Big) + \tfrac{1}{2}\,\mathrm{trace}\,\Big(\Sigma(t)^T \Sigma(t)\Big) \\
&= -\delta (x - x^*)^T \mathsf{L}(x - x^*) - (x - x^*)^T \Big(\nabla h(x) - \nabla h(x^*)\Big) + \tfrac{1}{2}|\Sigma(t)|_{\mathcal{F}}^2 \\
&\leq -\gamma(\|x - x^*\|_2^2) + \tfrac{1}{2}|\Sigma(t)|_{\mathcal{F}}^2.
\end{aligned}
$$

We note that $\mathrm{W} \in \mathcal{C}(\mathbb{R}^n; \mathbb{R}_{\geq 0})$ defined by $\mathrm{W}(x) \triangleq \gamma(\|x - x^*\|_2^2)$ verifies

$$
\mathrm{V}(x) = \tfrac{1}{2}\gamma^{-1}\big(\mathrm{W}(x)\big) \qquad \forall x \in \mathbb{R}^n,
$$

where γ^{-1} is concave and belongs to the class $\mathcal{K}_\infty$ as explained in Section 2.2. Therefore, V is a noise-dissipative Lyapunov function for (3.8), with concave $\eta \in \mathcal{K}_\infty$ given by $\eta(r) = 1/2\gamma^{-1}(r)$ and $\sigma \in \mathcal{K}$ given by $\sigma(r) \triangleq 1/2\, r^2$. •

The next result generalizes [4, Th. 4.1] to positive semidefinite Lyapunov functions that satisfy weaker dissipativity properties (cf. (3.5)) than the typical exponential-like inequality (3.6), and characterizes the overshoot gain.

Theorem 3.6. (Noise-dissipative Lyapunov functions have an NSS dynamics). *Under Assumption 2.1, and further assuming that Σ is continuous, suppose that* V *is a noise-dissipative Lyapunov function for* (2.1). *Then,*

$$
\mathbb{E}\big[\mathrm{V}(x(t))\big] \;\leq \tilde{\mu}\big(\mathrm{V}(x_0), t - t_0\big) + \eta\Big(2\,\sigma\big(\max_{t_0 \leq s \leq t} |\Sigma(s)|_{\mathcal{F}}\big)\Big), \tag{3.9}
$$

for all $t \geq t_0$, where the class $\mathcal{KL}$ function $(r,s) \mapsto \tilde{\mu}(r,s)$ is well defined as the solution $y(s)$ to the initial value problem

$$
\dot{y}(s) = -\tfrac{1}{2}\eta^{-1}(y(s)), \quad y(0) = r. \tag{3.10}
$$

Proof. Recall that Assumption 2.1 guarantees the global existence and uniqueness of solutions of (2.1). Given the process $\{\mathrm{V}(x(t))\}_{t \geq t_0}$, the proof strategy is to obtain a differential inequality for $\mathbb{E}\big[\mathrm{V}(x(t))\big]$ using Itô formula (2.3), and then use a comparison principle to translate the problem into one of standard input-to-state stability for an appropriate choice of the input.

To carry out this strategy, we consider Itô formula (2.3) with respect to an arbitrary reference time instant $t' \geq t_0$,

(3.11)

$$
\mathrm{V}(x(t)) = \mathrm{V}(x(t')) + \int_{t'}^{t} \mathcal{L}[\mathrm{V}](x(s),s)\mathrm{d}s + \int_{t'}^{t} \nabla \mathrm{V}(x(s))^T G(x(s),s)\Sigma(s)\mathrm{dB}(s),
$$

and we first ensure that the expectation of the integral against Brownian motion is 0. Let $S_k = \{x \in \mathbb{R}^n : \|x\|_2 \le k\}$ be the ball of radius k centered at the origin. Fix $x_0 \in \mathbb{R}^n$ and denote by τ_k the first exit time of $x(t)$ from S_k for integer values of k greater than $\|x(t_0)\|_2$, namely, $\tau_k \triangleq \inf\{s \ge t_0 : \|x(s)\|_2 \ge k\}$, for $k > \lceil \|x(t_0)\|_2 \rceil$. Since the event $\{\omega \in \Omega : \tau_k \le t\}$ belongs to $\mathcal{F}_t$ for each $t \ge t_0$, it follows that τ_k is an $\{\mathcal{F}_t\}$-stopping time for each $t \ge t_0$. Now, for each k fixed, if we consider the random variable $t_k(t) \triangleq \min\{t, \tau_k\}$ and define $I(t', t)$ as the stochastic integral in (3.11) for any fixed $t' \in [t_0, t_k(t)]$, then the process $I(t', t_k(t))$ has zero expectation as we show next. The function $X : S_k \times [t', t] \to \mathbb{R}$ given by $X(x, s) \triangleq \nabla \mathrm{V}(x)^T G(x, s)\Sigma(s)$ is essentially bounded (in its domain), and thus $\mathbb{E}\big[\int_{t'}^{t} \mathbf{1}_{[t', t_k(t)]}(s)\, X(x(s), t)^2 \,\mathrm{d}s\big] < \infty$, where $\mathbf{1}_{[t', t_k(t)]}(s)$ is the indicator function of the set $[t', t_k(t)]$. Therefore, $\mathbb{E}\big[I(t', t_k(t))\big] = 0$ by [14, Th. 5.16, p. 26]. Define now $\bar{\mathrm{V}}(t) \triangleq \mathbb{E}\big[\mathrm{V}(x(t))\big]$ and $\bar{\mathrm{W}}(t) \triangleq \mathbb{E}\big[\mathrm{W}(x(t))\big]$ in $\Gamma(t_0) \triangleq \{t \ge t_0 : \bar{\mathrm{V}}(t) < \infty\}$. By the above, taking expectations in (3.11) and using (3.4), we obtain that

$$
\begin{aligned}
\bar{\mathrm{V}}(t_k(t)) &= \bar{\mathrm{V}}(t') + \mathbb{E}\Big[\int_{t'}^{t_k(t)} \mathcal{L}[\mathrm{V}](x(s), s)\mathrm{d}s \Big] \\
&\le \bar{\mathrm{V}}(t') - \mathbb{E}\Big[\int_{t'}^{t_k(t)} \mathrm{W}(x(s))\mathrm{d}s \Big] + \mathbb{E}\Big[\int_{t'}^{t_k(t)} \sigma(|\Sigma(s)|_{\mathcal{F}})\, \mathrm{d}s \Big]
\end{aligned} \tag{3.12}
$$

for all $t \in \Gamma(t_0)$ and any $t' \in [t_0, t_k(t)]$. Next we use the fact that V is continuous and $\{x(t)\}_{t \ge t_0}$ is also continuous with probability 1. In addition, according to Fatou's lemma [16, p. 123] for convergence in the probability measure, we get that

$$
\begin{aligned}
\bar{\mathrm{V}}(t) &= \mathbb{E}\big[\mathrm{V}(x(\liminf_{k\to\infty}\, t_k(t))) \big] = \mathbb{E}\big[\liminf_{k\to\infty} \mathrm{V}(x(t_k(t))) \big] \\
&\le \liminf_{k\to\infty}\, \mathbb{E}\big[\mathrm{V}(x(t_k(t))) \big] = \liminf_{k\to\infty} \bar{\mathrm{V}}(t_k(t))
\end{aligned} \tag{3.13}
$$

for all $t \in \Gamma(t_0)$. Moreover, using the monotone convergence [16, p. 176] when $k \to \infty$ in both Lebesgue integrals in (3.12) (because both integrands are nonnegative and $\mathbf{1}_{[t', t_k(t)]}$ converges monotonically to $\mathbf{1}_{[t', t]}$ as $k \to \infty$ for any $t' \in [t_0, t_k(t)]$), we obtain from (3.13) that

$$
\bar{\mathrm{V}}(t) \le \bar{\mathrm{V}}(t') - \mathbb{E}\Big[\int_{t'}^{t} \mathrm{W}(x(s))\mathrm{d}s \Big] + \int_{t'}^{t} \sigma(|\Sigma(s)|_{\mathcal{F}})\, \mathrm{d}s \tag{3.14}
$$

for all $t \in \Gamma(t_0)$ and any $t' \in [t_0, t]$. Before resuming the argument we make two observations. First, applying Tonelli's theorem [16, p. 212] to the nonnegative process $\{\mathrm{W}(x(s))\}_{s \ge t'}$, it follows that

$$
\mathbb{E}\Big[\int_{t'}^{t} \mathrm{W}(x(s))\mathrm{d}s \Big] = \int_{t'}^{t} \bar{\mathrm{W}}(x(s))\mathrm{d}s. \tag{3.15}
$$

Second, using (3.3) and Jensen's inequality [1, Ch. 3], we get that

$$
\bar{\mathrm{V}}(t) = \mathbb{E}\big[\mathrm{V}(x(t))\big] \le \mathbb{E}\big[\eta(\mathrm{W}(x(t)))\big] \le \eta\big(\mathbb{E}\big[\mathrm{W}(x(t))\big]\big) = \eta\big(\bar{\mathrm{W}}(t)\big), \tag{3.16}
$$

because η is concave, so $\bar{\mathrm{W}}(t) \ge \eta^{-1}(\bar{\mathrm{V}}(t))$. Hence, (3.14) and (3.15) yield

$$
\begin{aligned}
\bar{\mathrm{V}}(t) &\le \bar{\mathrm{V}}(t') - \int_{t'}^{t} \bar{\mathrm{W}}(s)\, \mathrm{d}s + \int_{t'}^{t} \sigma(|\Sigma(s)|_{\mathcal{F}})\, \mathrm{d}s \\
&\le \bar{\mathrm{V}}(t') + \int_{t'}^{t} \Big(-\eta^{-1}(\bar{\mathrm{V}}(s)) + \sigma(|\Sigma(s)|_{\mathcal{F}}) \Big)\, \mathrm{d}s
\end{aligned} \tag{3.17}
$$

for all $t \in \Gamma(t_0)$ and any $t' \in [t_0, t]$, which in particular shows that $\Gamma(t_0)$ can be taken equal to $[t_0, \infty)$.

Now the strategy is to compare $\bar{\mathrm{V}}$ with the unique solution of an ordinary differential equation that represents an input-to-state stable (ISS) system. First we leverage the integral inequality (3.17) to show that $\bar{\mathrm{V}}$ is continuous in $[t_0, \infty)$, which allows us then to rewrite (3.17) as a differential inequality at t'. To to show that $\bar{\mathrm{V}}$ is continuous, we use the dominated convergence theorem [14, Thm. 2.3, P. 6] applied to $V_k(\hat{t}) \triangleq \mathrm{V}(x(\hat{t}))-\mathrm{V}(x(\hat{t}+1/k))$, for $\hat{t} \in [t_0, t]$, and similarly taking $\hat{t}-1/k$ (excluding, respectively, the cases when $\hat{t} = t$ or $\hat{t} = t_0$). The hypotheses are satisfied because V_k can be majorized using (3.17) as

$$|V_k(\hat{t})| \;\le\; \mathrm{V}(x(\hat{t})) + \mathrm{V}(x(\hat{t}+1/k)) \;\le\; 2\big(\mathrm{V}(x_0) + \int_{t_0}^{t} \sigma(|\Sigma(s)|_{\mathcal{F}})\, \mathrm{d}s\big), \tag{3.18}$$

where the term on the right is not a random variable and thus coincides with its expectation. Therefore, for every $\hat{t} \in [t_0, t]$,

$$\lim_{s\to\hat{t}} \mathbb{E}\big[\mathrm{V}(x(s))\big] = \mathbb{E}\big[\lim_{s\to\hat{t}} \mathrm{V}(x(s))\big] = \mathbb{E}\big[\mathrm{V}(x(\hat{t}))\big],$$

so $\bar{\mathrm{V}}$ is continuous on $[t_0, t]$, for any $t \ge t_0$. Now, using again (3.17) and the continuity of the integrand, we can bound the upper right-hand derivative [6, Appendix C.2] (also called upper Dini derivative), as

$$\begin{aligned} D^+\bar{\mathrm{V}}(t') &\triangleq \limsup_{t\to t',\, t>t'} \frac{\bar{\mathrm{V}}(t) - \bar{\mathrm{V}}(t')}{t-t'} \\ &\le \limsup_{t\to t',\, t>t'} \frac{1}{t-t'} \int_{t'}^{t} \big(-\eta^{-1}(\bar{\mathrm{V}}(s)) + \sigma(|\Sigma(s)|_{\mathcal{F}})\big)\, \mathrm{d}s \;=\; h(\bar{\mathrm{V}}(t'), d(t')), \end{aligned}$$

for any $t' \in [t_0, \infty)$, where the function $h : \mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0} \to \mathbb{R}$ is given by

$$h(y, d) \triangleq -\eta^{-1}(y) + d,$$

and $d(t) \triangleq \sigma(|\Sigma(t)|_{\mathcal{F}})$, which is continuous in $[t_0, \infty)$. Therefore, according to the comparison principle [6, Lemma 3.4, P. 102], using that $\bar{\mathrm{V}}$ is continuous in $[t_0, \infty)$ and $D^+\bar{\mathrm{V}}(t') \le h(\bar{\mathrm{V}}(t'), d(t'))$, for any $t' \in [t_0, \infty)$, the solutions [22, Sec. C.2] of the initial value problem

$$\dot{\mathrm{U}}(t) = h(\mathrm{U}(t), d(t)), \qquad \mathrm{U}_0 \triangleq \mathrm{U}(t_0) = \bar{\mathrm{V}}(t_0) \tag{3.19}$$

(where h is locally Lipschitz in the first argument as we show next), satisfy that $\mathrm{U}(t) \ge \bar{\mathrm{V}}(t)$ (≥ 0) in the common interval of existence. We argue the global existence and uniqueness of solutions of (3.19) as follows. Since $\alpha \triangleq \eta^{-1}$ is convex and class $\mathcal{K}_\infty$ (see Section 2.2), it holds that

$$\alpha(s') \le \alpha(s) \le \alpha(s') + \tfrac{\alpha(s'')-\alpha(s')}{s''-s'}(s-s')$$

for all $s \in [s', s'']$, for any $s'' > s' \ge 0$. Thus, $|\alpha(s)-\alpha(s')| = \alpha(s)-\alpha(s') \le L(s-s')$, for any $s'' \ge s \ge s' \ge 0$, where $L \triangleq (\alpha(s'') - \alpha(s'))/(s'' - s')$, so η^{-1} is locally Lipschitz. Hence, h is locally Lipschitz in $\mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0}$. Therefore, given the input function d and any $\mathrm{U}_0 \ge 0$, there is a unique maximal solution of (3.19), denoted

by $\mathrm{U}(\mathrm{U}_0, t_0; t)$, defined in a maximal interval $[t_0, t_{\max}(\mathrm{U}_0, t_0))$. (As a by-product, the initial value problem (3.10), which can be written as $\dot{y}(s) = \frac{1}{2} h(y(s), 0)$, $y(0) = r$, has a unique and strictly decreasing solution in $[0, \infty)$, so $\tilde{\mu}$ in the statement is well defined and in class $\mathcal{KL}$.) To show that (3.19) is ISS we follow a similar argument as in the proof of [23, Th. 5] (and note that, as a consequence, we obtain that $t_{\max}(\mathrm{U}_0, t_0) = \infty$). Firstly, if $\eta^{-1}(\mathrm{U}) \geq 2d$, then $\dot{\mathrm{U}}(t) = -\frac{1}{2}\eta^{-1}(\mathrm{U}(t))$, which implies that U is nonincreasing outside the set $S \triangleq \{t \geq t_0 : \mathrm{U}(t) \leq \eta(2d(t))\}$. Thus, if some $t^* \geq t_0$ belongs to S, then so does every $t \in [t^*, t_{\max}(\mathrm{U}_0, t_0))$ implying that U is locally bounded because d is locally bounded (in fact, continuous). (Note that $\mathrm{U}(t) \geq 0$ because $\dot{\mathrm{U}}(t) \geq 0$ whenever $\mathrm{U}(t) = 0$.) Therefore, for all $t \geq t_0$, and for $\tilde{\mu}$ as in the statement (which we have shown is well defined), we have that

$$\bar{\mathrm{V}}(t) \leq \mathrm{U}(t) \leq \max\Big\{\tilde{\mu}\big(\bar{\mathrm{V}}(t_0), t - t_0\big), \eta\Big(2 \max_{t_0 \leq s \leq t} d(s)\Big)\Big\}.$$

Since the maximum of two quantities is upper bounded by the sum, and using the definition of d together with the monotonicity of σ, it follows that

$$\bar{\mathrm{V}}(t) \leq \mathrm{U}(t) \leq \tilde{\mu}\big(\mathrm{V}(x_0), t - t_0\big) + \eta\Big(2\sigma\Big(\max_{t_0 \leq s \leq t} |\Sigma(s)|_{\mathcal{F}}\Big)\Big), \tag{3.20}$$

for all $t \geq t_0$, where we also used that $\bar{\mathrm{V}}(t_0) = \mathrm{V}(x_0)$, and the proof is complete. $\square$

Of particular interest to us is the case when the function V is lower and upper bounded by class $\mathcal{K}_\infty$ functions of the distance to a closed, not necessarily bounded, set.

Definition 3.7. (NSS-Lyapunov functions). *A function* $\mathrm{V} \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R}_{\geq 0})$ *is a* strong NSS-Lyapunov function in probability with respect to $\mathcal{U} \subseteq \mathbb{R}^n$ *for* (2.1) *if* V *is a noise-dissipative Lyapunov function and, in addition, there exist* $p > 0$ *and class* $\mathcal{K}_\infty$ *functions* α_1 *and* α_2 *such that*

$$\alpha_1(|x|_{\mathcal{U}}^p) \leq \mathrm{V}(x) \leq \alpha_2(|x|_{\mathcal{U}}^p), \quad \forall x \in \mathbb{R}^n. \tag{3.21}$$

If, moreover, α_1 *is convex, then* V *is a* pth moment NSS-Lyapunov function with respect to $\mathcal{U}$*.*

Note that a strong NSS-Lyapunov function in probability with respect to a set satisfies an inequality of the type (3.21) for any $p > 0$, whereas the choice of p is relevant when α_1 is required to be convex. The reason for the 'strong' terminology is that we require (3.5) to be satisfied with convex $\eta^{-1} \in \mathcal{K}_\infty$. Instead, a standard NSS-Lyapunov function in probability satisfies the same inequality with a class $\mathcal{K}_\infty$ function which is not necessarily convex. We also note that (3.21) implies that $\mathcal{U} = \{x \in \mathbb{R}^n : \mathrm{V}(x) = 0\}$, which is closed because V is continuous.

Example 3.8. (Example 3.5–revisited: an NSS-Lyapunov function). Consider the function V introduced in Example 3.5. For each $p \in (0, 2]$, note that

$$\alpha_{1p}(\|x - x^*\|_2^p) \leq \mathrm{V}(x) \leq \alpha_{2p}(\|x - x^*\|_2^p) \qquad \forall x \in \mathbb{R}^n,$$

for the convex functions $\alpha_{1p}(r) = \alpha_{2p}(r) \triangleq r^{2/p}$, which are in the class $\mathcal{K}_\infty$. (Recall that α_2 in Definition 3.7 is only required to be $\mathcal{K}_\infty$.) Thus, the function V is a pth moment NSS-Lyapunov function for (3.8) with respect to x^* for $p \in (0, 2]$. •

The notion of NSS-Lyapunov function plays a key role in establishing our main result on the stability of SDEs with persistent noise.

COROLLARY 3.9. (The existence of an NSS-Lyapunov function implies the corresponding NSS property). *Under Assumption 2.1, and further assuming that Σ is continuous, given a closed set $\mathcal{U} \subset \mathbb{R}^n$,*

(i) if $\mathrm{V} \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R}_{\geq 0})$ is a strong NSS-Lyapunov function in probability with respect to $\mathcal{U}$ for (2.1), then the system is NSS in probability with respect to $\mathcal{U}$ with gain functions

$$\mu(r,s) \triangleq \alpha_1^{-1}\big(\tfrac{2}{\epsilon}\tilde{\mu}(\alpha_2(r^p),s)\big), \quad \theta(r) \triangleq \alpha_1^{-1}\big(\tfrac{2}{\epsilon}\eta(2\sigma(r))\big); \tag{3.22}$$

(ii) if $\mathrm{V} \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R}_{\geq 0})$ is a pthNSS-Lyapunov function with respect to $\mathcal{U}$ for (2.1), then the system is pth moment NSS with respect to $\mathcal{U}$ with gain functions μ and θ as in (3.22) setting $\epsilon = 1$.

Proof. To show *(i)*, note that, since $\alpha_1(|x|_{\mathcal{U}}^p) \leq \mathrm{V}(x)$ for all $x \in \mathbb{R}^n$, with $\alpha_1 \in \mathcal{K}_\infty$, it follows that for any $\hat{\rho} > 0$ and $t \geq t_0$,

$$\begin{aligned}\mathbb{P}\Big\{|x(t)|_{\mathcal{U}}^p > \hat{\rho}\Big\} &= \mathbb{P}\Big\{\alpha_1(|x(t)|_{\mathcal{U}}^p) > \alpha_1(\hat{\rho})\Big\} \leq \mathbb{P}\Big\{\mathrm{V}(x(t)) > \alpha_1(\hat{\rho})\Big\} \leq \frac{\mathbb{E}\big[\mathrm{V}(x(t))\big]}{\alpha_1(\hat{\rho})} \\ &\leq \frac{1}{\alpha_1(\hat{\rho})}\bigg(\tilde{\mu}\Big(\alpha_2(|x_0|_{\mathcal{U}}^p),\, t-t_0\Big) + \eta\Big(2\sigma\big(\max_{t_0\leq s\leq t}|\Sigma(s)|_{\mathcal{F}}\big)\Big)\bigg),\end{aligned} \tag{3.23}$$

where we have used the strict monotonicity of α_1 in the first equation, Chebyshev's inequality [1, Ch. 3] in the second inequality, and the upper bound for $\mathbb{E}\big[\mathrm{V}(x(t))\big]$ obtained in Theorem 3.6, cf. (3.9), in the last inequality (leveraging the monotonicity of $\tilde{\mu}$ in the first argument and the fact that $\mathrm{V}(x) \leq \alpha_2(|x|_{\mathcal{U}}^p)$ for all $x \in \mathbb{R}^n$). Also, for any function $\alpha \in \mathcal{K}$, we have that $\alpha(2r) + \alpha(2s) \geq \alpha(r+s)$ for all $r, s \geq 0$. Thus,

$$\begin{aligned}\rho(\epsilon, x_0, t) &\triangleq \mu\big(|x_0|_{\mathcal{U}}, t-t_0\big) + \theta\Big(\max_{t_0\leq s\leq t}|\Sigma(s)|_{\mathcal{F}}\Big) \\ &\geq \alpha_1^{-1}\left(\frac{1}{\epsilon}\tilde{\mu}\Big(\alpha_2(|x_0|_{\mathcal{U}}^p), t-t_0\Big) + \frac{1}{\epsilon}\eta\Big(2\sigma\big(\max_{t_0\leq s\leq t}|\Sigma(s)|_{\mathcal{F}}\big)\Big)\right) \triangleq \hat{\rho}(\epsilon).\end{aligned} \tag{3.24}$$

Substituting now $\hat{\rho} \triangleq \hat{\rho}(\epsilon)$ in (3.23), and using that $\rho(\epsilon, x_0, t) \geq \hat{\rho}(\epsilon)$, we get that $\mathbb{P}\big\{|x(t)|_{\mathcal{U}}^p > \rho(\epsilon, x_0, t)\big\} \leq \mathbb{P}\big\{|x(t)|_{\mathcal{U}}^p > \hat{\rho}(\epsilon)\big\} \leq \epsilon$.

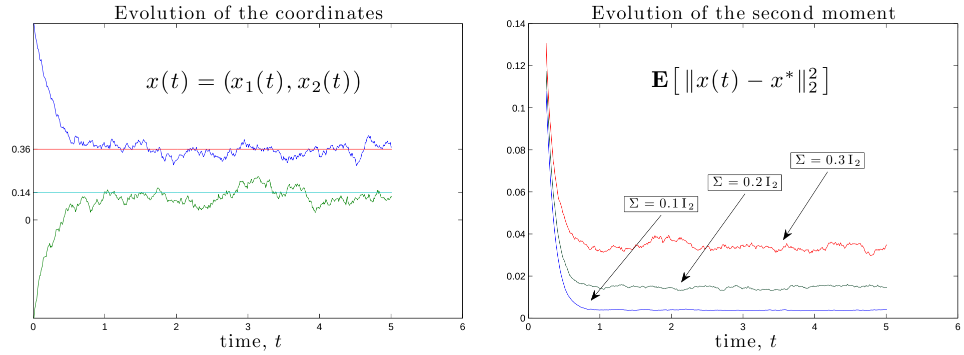
To show *(ii)*, since α_1^{-1} is concave, applying Jensen's inequality [1, Ch. 3], we get

$$\mathbb{E}\big[|x(t)|_{\mathcal{U}}^p\big] \leq \mathbb{E}\big[\alpha_1^{-1}\big(\mathrm{V}(x(t))\big)\big] \leq \alpha_1^{-1}\Big(\mathbb{E}\big[\mathrm{V}(x(t))\big]\Big) \leq \hat{\rho}(1) \leq \rho(1, x_0, t),$$

where in the last two inequalities we have used the bound for $\mathbb{E}\big[\mathrm{V}(x(t))\big]$ in (3.23) and the definition of $\hat{\rho}(\epsilon)$ in (3.24). ☐

EXAMPLE 3.10. (Example 3.5–revisited: illustration of Corollary 3.9). Consider again Example 3.5. Since V is a pth moment NSS-Lyapunov function for (3.8) with respect to the point x^* for $p \in (0, 2]$, as shown in Example 3.8, Corollary 3.9 implies that

$$\mathbb{E}\big[\|x - x^*\|_2^p\big] \leq \mu\big(\|x_0 - x^*\|_2, t-t_0\big) + \theta\Big(\max_{t_0\leq s\leq t}|\Sigma(s)|_{\mathcal{F}}\Big), \tag{3.25}$$



(a) (b)

Fig. 3.1: Evolution of the dynamics (3.8) with $\mathsf{L} = 0$, $h(x_1, x_2) = \log\big(e^{(x_1-2)} + e^{(x_2+1)}\big) + 0.5(x_1+x_2-1)^2 + (x_1-x_2)^2$, and initial condition $[x_1(0), x_2(0)] = (1, -0.5)$. Since h is a sum of convex functions, and the Hessian of the quadratic part of h has eigenvalues $\{2, 4\}$, we can take γ given by $\gamma(r) = 2\,r$, for $r \geq 0$. Plot (a) shows the evolution of the first and second coordinates with $\Sigma = 0.1\,\mathrm{I}_2$. Plot (b) illustrates the noise-to-state stability property in second moment with respect to $x^* = (0.36, 0.14)$, where the matrix $\Sigma(t)$ is a constant multiple of the identity. (The expectation is computed averaging over 500 realizations of the noise.)

for all $t \geq t_0$, $x_0 \in \mathbb{R}^n$, and $p \in (0, 2]$, where

$$\mu(r, s) \;=\; \big(2\tilde{\mu}(r^2, s)\big)^{p/2}, \quad \theta(r) \;=\; \big(\gamma^{-1}(r^2)\big)^{p/2},$$

and the class $\mathcal{KL}$ function $\tilde{\mu}$ is defined as the solution to the initial value problem (3.10) with $\eta(r) = \frac{1}{2}\gamma^{-1}(r)$. Figure 3.1 illustrates this noise-to-state stability property. We note that if the function h is strongly convex, i.e., if $\gamma(r) = c_\gamma\, r$ for some constant $c_\gamma > 0$, then $\tilde{\mu} : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ becomes $\tilde{\mu}(r, s) = re^{-c_\gamma s}$, and $\mu(r, s) = 2^{p/2}\, r^p e^{-c_\gamma p/2\, s}$, so the bound for $\mathbb{E}\big[\|x - x^*\|_2^p\big]$ in (3.25) decays exponentially with time to $\theta\big(\max_{t_0 \leq s \leq t} |\Sigma(s)|_{\mathcal{F}}\big)$. •

4. Refinements of the notion of proper functions. In this section, we analyze in detail the inequalities between functions that appear in the definition of noise-dissipative Lyapunov function, strong NSS-Lyapunov function in probability, and pth moment NSS-Lyapunov function. In Section 4.1, we establish that these inequalities can be regarded as equivalence relations. In Section 4.2, we make a complete characterization of the properties of two functions related by these equivalence relations. Finally, in Section 4.3, these results lead us to obtain an alternative formulation of Corollary 3.9.

4.1. Proper functions and equivalence relations. Here, we provide a refinement of the notion of proper functions with respect to each other. Proper functions

play an important role in stability analysis, see e.g., [6, 23].

DEFINITION 4.1. (Refinements of the notion of proper functions with respect to each other). *Let* $\mathcal{D} \subseteq \mathbb{R}^n$ *and the functions* $\mathrm{V}, \mathrm{W} : \mathcal{D} \to \mathbb{R}_{\geq 0}$ *be such that*

$$\alpha_1(\mathrm{W}(x)) \leq \mathrm{V}(x) \leq \alpha_2(\mathrm{W}(x)), \quad \forall x \in \mathcal{D},$$

for some functions $\alpha_1, \alpha_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. *Then,*

(i) if $\alpha_1, \alpha_2 \in \mathcal{K}$, *we say that* V *is* $\mathcal{K}$ - dominated by W *in* $\mathcal{D}$, *and write* $\mathrm{V} \lhd^{\mathcal{K}} \mathrm{W}$ *in* $\mathcal{D}$;

(ii) if $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$, *we say that* V *and* W *are* $\mathcal{K}_\infty$ - proper with respect to each other in $\mathcal{D}$, *and write* $\mathrm{V} \sim^{\mathcal{K}_\infty} \mathrm{W}$ *in* $\mathcal{D}$;

(iii) if $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ *are convex and concave, respectively, we say that* V *and* W *are* $\mathcal{K}_\infty^{cc}$ - (convex-concave) proper with respect to each other in $\mathcal{D}$, *and write* $\mathrm{V} \sim^{\mathcal{K}_\infty^{cc}} \mathrm{W}$ *in* $\mathcal{D}$;

(iv) if $\alpha_1(r) \triangleq c_{\alpha_1} r$ *and* $\alpha_2(r) \triangleq c_{\alpha_2} r$, *for some constants* $c_{\alpha_1}, c_{\alpha_2} > 0$, *we say that* V *and* W *are* equivalent in $\mathcal{D}$, *and write* $\mathrm{V} \sim \mathrm{W}$ *in* $\mathcal{D}$.

Note that the relations in Definition 4.1 are nested, i.e., given $\mathrm{V}, \mathrm{W} : \mathcal{D} \to \mathbb{R}_{\geq 0}$, the following chain of implications hold in $\mathcal{D}$:

$$\mathrm{V} \sim \mathrm{W} \;\Rightarrow\; \mathrm{V} \sim^{\mathcal{K}_\infty^{cc}} \mathrm{W} \;\Rightarrow\; \mathrm{V} \sim^{\mathcal{K}_\infty} \mathrm{W} \;\Rightarrow\; \mathrm{V} \lhd^{\mathcal{K}} \mathrm{W}. \tag{4.1}$$

Also, note that if $\mathrm{W}(x) = \|x\|_2$, $\mathcal{D}$ is a neighborhood of 0, and α_1, α_2 are class $\mathcal{K}$, then we recover the notion of V being a *proper* function [6]. If $\mathcal{D} = \mathbb{R}^n$, and V and W are seminorms, then the relation $\sim$ corresponds to the concept of equivalent seminorms.

The relation $\sim^{\mathcal{K}_\infty}$ is relevant for ISS and NSS in probability, whereas the relation $\sim^{\mathcal{K}_\infty^{cc}}$ is important for pth moment NSS. The latter is because the inequalities in $\sim^{\mathcal{K}_\infty^{cc}}$ are still valid, thanks to Jensen inequality, if we substitute V and W by their expectations along a stochastic process. Another fact about the relation $\sim^{\mathcal{K}_\infty^{cc}}$ is that $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$, convex and concave, respectively, must be asymptotically linear if $\mathrm{V}(\mathcal{D}) \supseteq [s_0, \infty)$, for some $s_0 \geq 0$, so that $\alpha_1(s) \leq \alpha_2(s)$ for all $s \geq s_0$. This follows from Lemma A.1.

REMARK 4.2. (Quadratic forms in a constrained domain). It is sometimes convenient to view the functions $\mathrm{V}, \mathrm{W} : \mathcal{D} \to \mathbb{R}_{\geq 0}$ as defined in a domain where their functional expression becomes simpler. To make this idea precise, assume there exist $i : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}^m$, with $m \geq n$, and $\hat{\mathrm{V}}, \hat{\mathrm{W}} : \hat{\mathcal{D}} \to \mathbb{R}_{\geq 0}$, where $\hat{\mathcal{D}} = i(\mathcal{D})$, such that $\mathrm{V} = \hat{\mathrm{V}} \circ i$ and $\mathrm{W} = \hat{\mathrm{W}} \circ i$. If this is the case, then the existence of $\alpha_1, \alpha_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\alpha_1\big(\hat{\mathrm{W}}(\hat{x})\big) \leq \hat{\mathrm{V}}(\hat{x}) \leq \alpha_2\big(\hat{\mathrm{W}}(\hat{x})\big)$, for all $\hat{x} \in \hat{\mathcal{D}}$, implies that $\alpha_1\big(\mathrm{W}(x)\big) \leq \mathrm{V}(x) \leq \alpha_2\big(\mathrm{W}(x)\big)$, for all $x \in \mathcal{D}$. The reason is that for any $x \in \mathcal{D}$ there exists $\hat{x} \in \hat{\mathcal{D}}$, given by $\hat{x} = i(x)$, such that $\mathrm{V}(x) = \hat{\mathrm{V}}(\hat{x})$ and $\mathrm{W}(x) = \hat{\mathrm{W}}(\hat{x})$, so

$$\alpha_1\big(\mathrm{W}(x)\big) = \alpha_1\big(\hat{\mathrm{W}}(\hat{x})\big) \leq \mathrm{V}(x) = \hat{\mathrm{V}}(\hat{x}) \leq \alpha_2\big(\hat{\mathrm{W}}(\hat{x})\big) = \alpha_2\big(\mathrm{W}(x)\big).$$

Consequently, if any of the relations given in Definition 4.1 is satisfied by $\hat{\mathrm{V}}$, $\hat{\mathrm{W}}$ in $\hat{\mathcal{D}}$, then the corresponding relation is satisfied by V, W in $\mathcal{D}$. For instance, in some scenarios this procedure can allow us to rewrite the original functions V, W as quadratic forms $\hat{\mathrm{V}}$, $\hat{\mathrm{W}}$ in a constrained set of an extended Euclidean space, where it is easier to establish the appropriate relation between the functions. We make use of this observation in Section 4.3 below. •

LEMMA 4.3. (Powers of seminorms with the same nullspace). *Let* A *and* B *in* $\mathbb{R}^{m \times n}$ *be nonzero matrices with the same nullspace,* $\mathcal{N}(A) = \mathcal{N}(B)$. *Then, for any* $p, q > 0$,

the inequalities $\alpha_1\big(\|x\|_A^p\big) \leq \|x\|_B^q \leq \alpha_2\big(\|x\|_A^p\big)$ *are verified with*

$$\alpha_1(r) \triangleq \Big(\frac{\lambda_{n-k}(B^TB)}{\lambda_{max}(A^TA)}\Big)^{\frac{q}{2}} r^{q/p}; \quad \alpha_2(r) \triangleq \Big(\frac{\lambda_{max}(B^TB)}{\lambda_{n-k}(A^TA)}\Big)^{\frac{q}{2}} r^{q/p},$$

where $k \triangleq \dim(\mathcal{N}(A))$. *In particular,* $\|.\|_A^p \sim \|.\|_B^p$ *and* $\|.\|_A^p \sim^{\mathcal{K}_\infty} \|.\|_B^q$ *in* $\mathbb{R}^n$ *for any real numbers* $p, q > 0$.

Proof. For $\mathcal{U} \triangleq \mathcal{N}(A)$, write any $x \in \mathbb{R}^n$ as $x = x_{\mathcal{U}} + x_{\mathcal{U}^\perp}$, where $x_{\mathcal{U}} \in \mathcal{U}$ and $x_{\mathcal{U}^\perp} \in \{x \in \mathbb{R}^n \,:\, x^Tu = 0 \,, \forall u \in \mathcal{U}\}$, so that $Ax = A(x_{\mathcal{U}} + x_{\mathcal{U}^\perp}) = Ax_{\mathcal{U}^\perp}$ and $Bx = Bx_{\mathcal{U}^\perp}$ because $\mathcal{N}(A) = \mathcal{N}(B) = \mathcal{U}$. Using the formulas for the eigenvalues in [5, p. 178], we see that the next chain of inequalities hold:

$$\begin{aligned}
&\alpha_1\big(\|x\|_A^p\big) = \alpha_1\Big(\big(x_{\mathcal{U}^\perp}^T A^TAx_{\mathcal{U}^\perp}\big)^{\frac{p}{2}}\Big) \leq \alpha_1\Big(\big(\lambda_{\max}(A^TA)x_{\mathcal{U}^\perp}^T x_{\mathcal{U}^\perp}\big)^{\frac{p}{2}}\Big) \\
&\leq \big(\lambda_{n-k}(B^TB)x_{\mathcal{U}^\perp}^T x_{\mathcal{U}^\perp}\big)^{\frac{q}{2}} \leq \big(x_{\mathcal{U}^\perp}^T B^TBx_{\mathcal{U}^\perp}\big)^{\frac{q}{2}} \leq \big(\lambda_{\max}(B^TB)x_{\mathcal{U}^\perp}^T x_{\mathcal{U}^\perp}\big)^{\frac{q}{2}} \\
&\leq \alpha_2\Big(\big(\lambda_{n-k}(A^TA)x_{\mathcal{U}^\perp}^T x_{\mathcal{U}^\perp}\big)^{\frac{p}{2}}\Big) \leq \alpha_2\Big(\big(x_{\mathcal{U}^\perp}^T A^TAx_{\mathcal{U}^\perp}\big)^{\frac{p}{2}}\Big) = \alpha_2\big(\|x\|_A^p\big),
\end{aligned}$$

where $\|x\|_B^q = \big(x_{\mathcal{U}^\perp}^T B^TBx_{\mathcal{U}^\perp}\big)^{\frac{q}{2}}$. From this we conclude that $\|.\|_A^p \sim^{\mathcal{K}_\infty} \|.\|_B^q$ in $\mathbb{R}^n$. Finally, when $p = q$, the class $\mathcal{K}_\infty$ functions α_1, α_2 in the statement are linear, so we obtain that $\|.\|_A^p \sim \|.\|_B^p$ in $\mathbb{R}^n$. ∎

Next we show that $\sim^{\mathcal{K}_\infty}$ and $\sim^{\mathcal{K}_\infty^{cc}}$ are reflexive, symmetric, and transitive, and hence define equivalence relations.

Lemma 4.4. (The $\mathcal{K}_\infty$- and $\mathcal{K}_\infty^{cc}$-proper relations are equivalence relations). *The relations* $\sim^{\mathcal{K}_\infty}$ *and* $\sim^{\mathcal{K}_\infty^{cc}}$ *in any set* $\mathcal{D} \subseteq \mathbb{R}^n$ *are both equivalence relations.*

Proof. For convenience, we represent both relations by $\sim^*$. Both are reflexive, i.e., $\mathrm{V} \sim^* \mathrm{V}$, because one can take $\alpha_1(r) = \alpha_2(r) = r$ noting that a linear function is both convex and concave. Both are symmetric, i.e., $\mathrm{V} \sim^* \mathrm{W}$ if and only if $\mathrm{W} \sim^* \mathrm{V}$, because if $\alpha_1 \circ \mathrm{W} \leq \mathrm{V} \leq \alpha_2 \circ \mathrm{W}$ in $\mathcal{D}$, then $\alpha_2^{-1} \circ \mathrm{V} \leq \mathrm{W} \leq \alpha_1^{-1} \circ \mathrm{V}$ in $\mathcal{D}$. In the case of $\sim^{\mathcal{K}_\infty}$, the inverse of a class $\mathcal{K}_\infty$ function is class $\mathcal{K}_\infty$. Additionally, in the case of $\sim^{\mathcal{K}_\infty^{cc}}$, if $\alpha \in \mathcal{K}_\infty$ is convex (respectively, concave), then $\alpha^{-1} \in \mathcal{K}_\infty$ is concave (respectively, convex). Finally, both are transitive, i.e., $\mathrm{U} \sim^* \mathrm{V}$ and $\mathrm{V} \sim^* \mathrm{W}$ imply $\mathrm{U} \sim^* \mathrm{W}$, because if $\alpha_1 \circ \mathrm{V} \leq \mathrm{U} \leq \alpha_2 \circ \mathrm{V}$ and $\tilde{\alpha}_1 \circ \mathrm{W} \leq \mathrm{V} \leq \tilde{\alpha}_2 \circ \mathrm{W}$ in $\mathcal{D}$, then $\alpha_1 \circ \tilde{\alpha}_1 \circ \mathrm{W} \leq \mathrm{U} \leq \alpha_2 \circ \tilde{\alpha}_2 \circ \mathrm{W}$ in $\mathcal{D}$. In the case of $\sim^{\mathcal{K}_\infty}$, the composition of two class $\mathcal{K}_\infty$ functions is class $\mathcal{K}_\infty$. Additionally, in the case of $\sim^{\mathcal{K}_\infty^{cc}}$, if $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ are both convex (respectively, concave), then the compositions $\alpha_1 \circ \alpha_2$ and $\alpha_2 \circ \alpha_1$ belong to $\mathcal{K}_\infty$ and are convex (respectively, concave), as explained in Section 2.2. ∎

Remark 4.5. (The relation $\lhd^{\mathcal{K}}$ is not an equivalence relation). The proof above also shows that the relation $\lhd^{\mathcal{K}}$ is reflexive and transitive. However, it is not symmetric: consider $\mathrm{V}, \mathrm{W} \in \mathcal{C}(\mathbb{R}^n; \mathbb{R}_{\geq 0})$ given by $\mathrm{V}(x) = 1 - e^{-\|x\|_2}$ and $\mathrm{W}(x) = \|x\|_2$. Clearly, $\mathrm{V} \lhd^{\mathcal{K}} \mathrm{W}$ in $\mathbb{R}^n$ by taking $\alpha_1 = \alpha_2 = \alpha \in \mathcal{K}$, with $\alpha(s) = 1 - e^{-s}$. On the other hand, if there exist $\tilde{\alpha}_1, \tilde{\alpha}_2 \in \mathcal{K}$ such that $\tilde{\alpha}_1(\mathrm{V}(x)) \leq \mathrm{W}(x) \leq \tilde{\alpha}_2(\mathrm{V}(x))$ for all $x \in \mathbb{R}^n$, then we reach the contradiction, by continuity of $\tilde{\alpha}_2$, that $\lim_{\|x\|_2 \to \infty} \|x\|_2 \leq \tilde{\alpha}_2\big(\lim_{\|x\|_2 \to \infty} \big(1 - e^{-\|x\|_2}\big)\big) = \tilde{\alpha}_2(1) < \infty$. •

4.2. Characterization of proper functions with respect to each other. In this section, we provide a complete characterization of the properties that two

functions must satisfy to be related by the equivalence relations defined in Section 4.1. For $\mathcal{D} \subseteq \mathbb{R}^n$, consider $\mathrm{V}_1, \mathrm{V}_2 : \mathcal{D} \to \mathbb{R}_{\geq 0}$. Given a real number $p > 0$, define

$$\phi_p(s) \triangleq \sup_{\{y \in \mathcal{D} \,:\, \mathrm{V}_2(y) \leq \sqrt[p]{s}\}} \mathrm{V}_1(y),$$
$$\psi_p(s) \triangleq \inf_{\{y \in \mathcal{D} \,:\, \mathrm{V}_2(y) \geq \sqrt[p]{s}\}} \mathrm{V}_1(y),$$

for $s \geq 0$. The value $\phi_p(s)$ gives the supremum of the function V_1 in the $\sqrt[p]{s}$ - sublevel set of V_2, and $\psi_p(s)$ is the infimum of V_1 in the $\sqrt[p]{s}$ - superlevel set of V_2. Thus, the functions ϕ_p and ψ_p satisfy

$$\psi_p\big(\mathrm{V}_2(x)^p\big) = \inf_{\substack{\{y \in \mathcal{D}: \\ \mathrm{V}_2(y) \geq \mathrm{V}_2(x)\}}} \mathrm{V}_1(y) \leq \mathrm{V}_1(x) \leq \sup_{\substack{\{y \in \mathcal{D}: \\ \mathrm{V}_2(y) \leq \mathrm{V}_2(x)\}}} \mathrm{V}_1(y) = \phi_p\big(\mathrm{V}_2(x)^p\big), \tag{4.2}$$

for all $x \in \mathcal{D}$, which suggests ϕ_p and ψ_p as pre-comparison functions to construct α_1 and α_2 in Definition 4.1. To this end, we find it useful to formulate the following properties of the function V_1 with respect to V_2:

P0: The set $\{x \in \mathcal{D} : \mathrm{V}_2(x) = s\}$ is nonempty for all $s \geq 0$.

P1: The nullsets of V_1 and V_2 are the same, i.e., $\{x \in \mathcal{D} : \mathrm{V}_1(x) = 0\} = \{x \in \mathcal{D} : \mathrm{V}_2(x) = 0\}$.

P2: The function ϕ_1 is locally bounded in $\mathbb{R}_{\geq 0}$ and right continuous at 0, and ψ_1 is positive definite.

P3: The next limit holds: $\lim_{s \to \infty} \psi_1(s) = \infty$.

P4 (as a function of $p > 0$): The asymptotic behavior of ϕ_p and ψ_p is such that $\phi_p(s)$ and $s^2/\psi_p(s)$ are both in $\mathcal{O}(s)$ as $s \to \infty$.

The next result shows that these properties completely characterize whether the functions V_1 and V_2 are related through the equivalence relations defined in Section 4.1. This result generalizes [6, Lemma 4.3] in several ways: the notions of proper functions considered here are more general and are not necessarily restricted to a relationship between an arbitrary function and the distance to a compact set.

Theorem 4.6. (Characterizations of proper functions with respect to each other). *Let* $\mathrm{V}_1, \mathrm{V}_2 : \mathcal{D} \to \mathbb{R}_{\geq 0}$*, and assume* V_2 *satisfies* P0*. Then*

(i) V_1 *satisfies* $\{\mathrm{P}i\}_{i=1}^2$ *with respect to* V_2 $\Leftrightarrow$ $\mathrm{V}_1 \lhd^{\mathcal{K}} \mathrm{V}_2$ *in* $\mathcal{D}$*;*
(ii) V_1 *satisfies* $\{\mathrm{P}i\}_{i=1}^3$ *with respect to* V_2 $\Leftrightarrow$ $\mathrm{V}_1 \sim^{\mathcal{K}_\infty} \mathrm{V}_2$ *in* $\mathcal{D}$*;*
(iii) V_1 *satisfies* $\{\mathrm{P}i\}_{i=1}^4$ *with respect to* V_2 *for* $p > 0$ $\Leftrightarrow$ $\mathrm{V}_1 \sim^{\mathcal{K}_\infty^{cc}} \mathrm{V}_2^p$ *in* $\mathcal{D}$*.*

Proof. We begin by establishing a few basic facts about the pre-comparison functions ψ_p and ϕ_p. By definition and by P0, it follows that $0 \leq \psi_1(s) \leq \phi_1(s)$ for all $s \geq 0$. Since ϕ_1 is locally bounded by P2, then so is ψ_1. In particular, ϕ_1 and ψ_1 are well defined in $\mathbb{R}_{\geq 0}$. Moreover, both ϕ_1 and ψ_1 are nondecreasing because if $s_2 \geq s_1$, then the supremum is taken in a larger set, $\{x \in \mathcal{D} : \mathrm{V}_2(x) \leq s_2\} \supseteq \{x \in \mathcal{D} : \mathrm{V}_2(x) \leq s_1\}$, and the infimum is taken in a smaller set, $\{x \in \mathcal{D} : \mathrm{V}_2(x) \geq s_2\} \subseteq \{x \in \mathcal{D} : \mathrm{V}_2(x) \geq s_1\}$. Furthermore, for any $q > 0$, the functions ϕ_q and ψ_q are also monotonic and positive definite because $\phi_q(s) = \phi_1(\sqrt[q]{s})$ and $\psi_q(s) = \psi_1(\sqrt[q]{s})$ for all $s \geq 0$. We now use these properties of the pre-comparison functions to construct α_1, α_2 in Definition 4.1 required by the implications from left to right in each statement.

Proof of *(i)* ($\Rightarrow$). To show the existence of $\alpha_2 \in \mathcal{K}$ such that $\alpha_2(s) \geq \phi_1(s)$ for all $s \in \mathbb{R}_{\geq 0}$, we proceed as follows. Since ϕ_1 is locally bounded and nondecreasing, given a strictly increasing sequence $\{b_k\}_{k\geq 1} \subseteq \mathbb{R}_{\geq 0}$ with $\lim_{k\to\infty} b_k = \infty$, we choose the sequence $\{M_k\}_{k\geq 1} \subseteq \mathbb{R}_{\geq 0}$, setting $M_0 = 0$, in the following way:

$$M_k \triangleq \max\Big\{ \sup_{s\in[0,b_k]} \phi_1(s)\,,\ M_{k-1} + 1/k^2 \Big\} = \max\big\{\phi_1(b_k)\,,\ M_{k-1} + 1/k^2\big\}. \tag{4.3}$$

This choice guarantees that $\{M_k\}_{k\geq 1}$ is strictly increasing and, for each $k \geq 1$,

$$0 \leq M_k - \phi_1(b_k) \leq \sum_{i=1}^{k} \frac{1}{i^2} \leq \pi^2/6. \tag{4.4}$$

Also, since ϕ_1 is right continuous at 0, we can choose $b_1 > 0$ such that there exists $\alpha_2 : [0, b_1] \to \mathbb{R}_{\geq 0}$ continuous, positive definite and strictly increasing, satisfying that $\alpha_2(s) \geq \phi_1(s)$ for all $s \in [0, b_1]$ and with $\alpha_2(b_1) = M_2$. (This is possible because the only function that cannot be upper bounded by an arbitrary continuous function in some arbitrarily small interval $[0, b_1]$ is the function that has a jump at 0.) The rest of the construction is explicit. We define α_2 as a piecewise linear function in (b_1, ∞) in the following way: for each $k \geq 2$, we define

$$\alpha_2(s) \triangleq \alpha_2(b_{k-1}) + \frac{M_{k+1} - \alpha_2(b_{k-1})}{b_k - b_{k-1}}\,(s - b_{k-1}), \qquad \forall s \in (b_{k-1}, b_k].$$

The resulting α_2 is continuous by construction. Also, $\alpha_2(b_1) = M_2$, so that, inductively, $\alpha_2(b_{k-1}) = M_k$ for $k \geq 2$. Two facts now follow: first, $M_{k+1} - \alpha_2(b_{k-1}) = M_{k+1} - M_k \geq 1/(k+1)^2$ for $k \geq 2$, so α_2 has positive slope in each interval $(b_{k-1}, b_k]$ and thus is strictly increasing in (b_1, ∞); second, $\alpha_2(s) > \alpha_2(b_{k-1}) = M_k \geq \phi_1(b_k) \geq \phi_1(s)$ for all $s \in (b_{k-1}, b_k]$, for each $k \geq 2$, so $\alpha_2(s) \geq \phi_1(s)$ for all $s \in (b_1, \infty)$.

We have left to show the existence of $\alpha_1 \in \mathcal{K}$ such that $\alpha_1(s) \leq \psi_1(s)$ for all $s \in \mathbb{R}_{\geq 0}$. First, since $0 \leq \psi_1(s) \leq \phi_1(s)$ for all $s \geq 0$ by definition and by P0, using the sandwich theorem [11, p. 107], we derive that ψ_1 is right continuous at 0 the same as ϕ_1. In addition, since ψ_1 is nondecreasing, it can only have a countable number of jump discontinuities (none of them at 0). Therefore, we can pick $c_1 > 0$ such that a continuous and nondecreasing function $\hat{\psi}_1$ can be constructed in $[0, c_1)$ by removing the jumps of ψ_1, so that $\hat{\psi}_1(s) \leq \psi_1(s)$. Moreover, since ψ_1 is positive definite and right continuous at 0, then $\hat{\psi}_1$ is also positive definite. Thus, there exists α_1 in $[0, c_1)$ continuous, positive definite, and strictly increasing, such that, for some $r < 1$,

$$\alpha_1(s) \leq r\hat{\psi}_1(s) \leq r\psi_1(s) \tag{4.5}$$

for all $s \in [0, c_1)$. To extend α_1 to a function in class $\mathcal{K}$ in $\mathbb{R}_{\geq 0}$, we follow a similar strategy as for α_2. Given a strictly increasing sequence $\{c_k\}_{k\geq 2} \subseteq \mathbb{R}_{\geq 0}$ with $\lim_{k\to\infty} c_k = \infty$, we define a sequence $\{m_k\}_{k\geq 1} \subseteq \mathbb{R}_{\geq 0}$ in the following way:

$$m_k \triangleq \inf_{s\in[c_k, c_{k+1})} \psi_1(s) - \tfrac{\psi_1(c_1) - \alpha_1(c_1)}{1+k^2} = \psi_1(c_k) - \tfrac{\psi_1(c_1) - \alpha_1(c_1)}{1+k^2}. \tag{4.6}$$

Next we define α_1 in $[c_1, \infty)$ as the piecewise linear function

$$\alpha_1(s) \triangleq \alpha_1(c_k) + \frac{m_k - \alpha_1(c_k)}{c_{k+1} - c_k}\,(s - c_k), \qquad \forall s \in [c_k, c_{k+1}),$$

for all $k \geq 1$, so α_1 is continuous by construction. It is also strictly increasing because $\alpha_1(c_2) = m_1 = (\psi_1(c_1) + \alpha_1(c_1))/2 > \alpha_1(c_1)$ by (4.5), and also, for each $k \geq 2$, the slopes are positive because $m_k - \alpha_1(c_k) = m_k - m_{k-1} > 0$ (due to the fact that $\{m_k\}_{k\geq 1}$ in (4.6) is strictly increasing because ψ_1 is nondecreasing). Finally, $\alpha_1(s) < \alpha_1(c_{k+1}) = m_k < \psi_1(c_k) \leq \psi_1(s)$ for all $s \in [c_k, c_{k+1})$, for all $k \geq 1$ by (4.6).

Equipped with α_1, α_2 as defined above, and as a consequence of (4.2), we have that

$$\alpha_1(\mathrm{V}_2(x)) \leq \psi_1(\mathrm{V}_2(x)) \leq \mathrm{V}_1(x) \leq \phi_1(\mathrm{V}_2(x)) \leq \alpha_2(\mathrm{V}_2(x)), \quad \forall x \in \mathcal{D}. \tag{4.7}$$

This concludes the proof of *(i)* ($\Rightarrow$).

As a preparation for *(ii)*-*(iii)* ($\Rightarrow$), and assuming P3, we derive two facts regarding the functions α_1 and α_2 constructed above. First, we establish that

$$\alpha_2(s) \in \mathcal{O}(\phi_1(s)) \text{ as } s \to \infty. \tag{4.8}$$

To show this, we argue that

$$\lim_{k\to\infty} \sup_{s\in(b_{k-1},b_k]} \big(\alpha_2(s) - \phi_1(s)\big) \leq \lim_{k\to\infty} \big(\phi_1(b_{k+1}) - \phi_1(b_{k-1})\big) + \pi^2/6, \tag{4.9}$$

so that there exist $C, s_1 > 0$ such that $\alpha_2(s) \leq 3\phi_1(s) + C$, for all $s \geq s_1$. Thus, noting that $\lim_{s\to\infty} \phi_1(s) = \infty$ as a consequence of P3, the expression (4.8) follows. To establish (4.9), we use the monotonicity of α_2 and ϕ_1, (4.3) and (4.4). For $k \geq 2$,

$$\begin{aligned}
&\sup_{s\in(b_{k-1},b_k]} \big(\alpha_2(s) - \phi_1(s)\big) \leq \alpha_2(b_k) - \phi_1(b_{k-1}) = M_{k+1} - \phi_1(b_{k-1}) \\
&= \max\big\{\phi_1(b_{k+1}) - \phi_1(b_{k-1})\,,\; M_k + 1/(k+1)^2 - \phi_1(b_{k-1})\big\} \\
&\leq \max\big\{\phi_1(b_{k+1}) - \phi_1(b_{k-1})\,,\; \phi_1(b_k) + \pi^2/6 + 1/(k+1)^2 - \phi_1(b_{k-1})\big\}.
\end{aligned}$$

Second, the construction of α_1 guarantees that

$$\psi_1(s) \in \mathcal{O}(\alpha_1(s)) \text{ as } s \to \infty, \tag{4.10}$$

because, as we show next,

$$\lim_{k\to\infty} \sup_{s\in[c_k,c_{k+1})} \big(\psi_1(s) - \alpha_1(s)\big) \leq \lim_{k\to\infty} \big(\alpha_1(c_{k+2}) - \alpha_1(c_k)\big), \tag{4.11}$$

so there exists $s_2 > 0$ such that $\psi_1(s) \leq 3\alpha_1(s)$ for all $s \geq s_2$. To obtain (4.11), we leverage the monotonicity of ψ_1 and α_1, and (4.6); namely, for $k \geq 2$,

$$\begin{aligned}
&\sup_{s\in[c_k,c_{k+1})} \big(\psi_1(s) - \alpha_1(s)\big) \leq \psi_1(c_{k+1}) - \alpha_1(c_k) \\
&= m_{k+1} + \tfrac{\psi_1(c_1)-\alpha_1(c_1)}{1+(k+1)^2} - \alpha_1(c_k) = \alpha_1(c_{k+2}) + \tfrac{\psi_1(c_1)-\alpha_1(c_1)}{1+(k+1)^2} - \alpha_1(c_k).
\end{aligned}$$

Equipped with (4.8) and (4.10), we prove next *(ii)*-*(iii)* ($\Rightarrow$).

Proof of *(ii)* ($\Rightarrow$): If, in addition, P3 holds, then $\lim_{s\to\infty} \phi_1(s) \geq \lim_{s\to\infty} \psi_1(s) = \infty$. This guarantees that $\alpha_2 \in \mathcal{K}_\infty$. Also, according to (4.10), P3 implies that α_1 is unbounded, and thus in $\mathcal{K}_\infty$ as well. The result now follows by (4.7).

Proof of *(iii)* ($\Rightarrow$): Finally, assume that P4 also holds for some $p > 0$. We show next the existence of the required convex and concave functions involved in the relation $\sim^{\mathcal{K}_\infty^{cc}}$. Let $\alpha_{1,p}(s) \triangleq \alpha_1(\sqrt[p]{s})$ and $\alpha_{2,p}(s) \triangleq \alpha_2(\sqrt[p]{s})$ for $s \geq 0$, so that

$$\alpha_{1,p}(s) = \alpha_1(\sqrt[p]{s}) \leq \psi_1(\sqrt[p]{s}) = \psi_p(s) \quad \text{and} \quad \phi_p(s) = \phi_1(\sqrt[p]{s}) \leq \alpha_2(\sqrt[p]{s}) = \alpha_{2,p}(s).$$

From (4.8) and P4, it follows that there exist s', c_1, $c_2 > 0$ such that $\alpha_2(s) \leq c_1\phi_1(s)$ and $\phi_p(s) \leq c_2 s$ for all $s \geq s'$. Thus,

$$\alpha_{2,p}(s) = \alpha_2(\sqrt[p]{s}) \leq c_1\phi_1(\sqrt[p]{s}) = c_1\phi_p(s) \leq c_1 c_2 s,$$

for all $s \geq s'$, so $\alpha_{2,p}(s)$ is in $\mathcal{O}(s)$ as $s \to \infty$. Similarly, according to (4.10) and P4, there are constants s'', c_3, $c_4 > 0$ such that $\psi_1(s) \leq c_3\,\alpha_1(s)$ and $s^2 \leq c_4 s\,\psi_p(s)$ for all $s \geq s''$. Thus,

$$s\,\alpha_{1,p}(s) = s\,\alpha_1(\sqrt[p]{s}) \geq s\,\tfrac{1}{c_3}\psi_1(\sqrt[p]{s}) = s\,\tfrac{1}{c_3}\psi_p(s) \geq \tfrac{1}{c_3 c_4}s^2,$$

for all $s \geq s''$, so $s^2/\alpha_{1,p}(s)$ is in $\mathcal{O}(s)$ as $s \to \infty$. Summarizing, the construction of α_1, α_2 guarantees, under P4, that $\alpha_{1,p}$, $\alpha_{2,p}$ satisfy that $s^2/\alpha_{1,p}(s)$ and $\alpha_{2,p}(s)$ are in $\mathcal{O}(s)$ as $s \to \infty$ (and, as a consequence, so are $s^2/\alpha_{2,p}(s)$ and $\alpha_{1,p}(s)$). Therefore, according to Lemma A.1, we can leverage (4.7) by taking $\tilde{\alpha}_1$, $\tilde{\alpha}_2 \in \mathcal{K}_\infty$, convex and concave, respectively, such that, for all $x \in \mathcal{D}$,

$$\begin{aligned}\tilde{\alpha}_1\big(\mathrm{V}_2(x)^p\big) &\leq \alpha_{1,p}(\mathrm{V}_2(x)^p) = \alpha_1(\mathrm{V}_2(x)) \leq \psi_1\big(\mathrm{V}_2(x)\big) \leq \mathrm{V}_1(x)\\ &\leq \phi_1\big(\mathrm{V}_2(x)\big) \leq \alpha_2(\mathrm{V}_2(x)) = \alpha_{2,p}(\mathrm{V}_2(x)^p) \leq \tilde{\alpha}_2\big(\mathrm{V}_2(x)^p\big).\end{aligned}$$

Proof of *(i)* ($\Leftarrow$): If there exist class $\mathcal{K}$ functions α_1, α_2 such that $\alpha_1(\mathrm{V}_2(x)) \leq \mathrm{V}_1(x) \leq \alpha_2(\mathrm{V}_2(x))$ for all $x \in \mathcal{D}$, then the nullsets of V_1 and V_2 are the same, which is the property P1. In addition, $0 \leq \phi_1(s) \leq \alpha_2(s)$ for all $s \geq 0$, so ϕ_1 is locally bounded and, moreover, the sandwich theorem guarantees that ϕ_1 is right continuous at 0. Also, since $\alpha_1(s) \leq \psi_1(s)$, for all $s \geq 0$, and $\psi_1(0) = 0$, it follows that ψ_1 is positive definite. Therefore, P2 also holds.

Proof of *(ii)* ($\Leftarrow$): Since $\psi_1(s) \geq \alpha_1(s)$ for all $s \geq 0$, the property P3 follows because

$$\lim_{s\to\infty} \psi_1(s) \geq \lim_{s\to\infty} \alpha_1(s) = \infty.$$

Proof of *(iii)* ($\Leftarrow$): If $\mathrm{V}_1 \sim^{\mathcal{K}_\infty^{cc}} \mathrm{V}_2^p$, then $\mathrm{V}_1 \sim^{\mathcal{K}_\infty} \mathrm{V}_2^p$ by (4.1). Also, we have trivially that $\mathrm{V}_2^p \sim^{\mathcal{K}_\infty} \mathrm{V}_2$. Since $\sim^{\mathcal{K}_\infty}$ is an equivalence relation by Lemma 4.4, it follows that $\mathrm{V}_1 \sim^{\mathcal{K}_\infty} \mathrm{V}_2$, so the properties $\{\mathrm{P}i\}_{i=1}^3$ hold as in *(ii)* ($\Leftarrow$). We have left to derive P4. If $\mathrm{V}_1 \sim^{\mathcal{K}_\infty^{cc}} \mathrm{V}_2^p$, then there exist $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ convex and concave, respectively, such that $\alpha_1\big(\mathrm{V}_2(x)^p\big) \leq \mathrm{V}_1(x) \leq \alpha_2\big(\mathrm{V}_2(x)^p\big)$ for all $x \in \mathcal{D}$. Hence, by the definition of ψ_p and ϕ_p, and P0, and by the monotonicity of α_1 and α_2, we have that, for all $s \geq 0$,

$$\begin{aligned}\alpha_1(s) &\leq \inf_{\{x\in\mathcal{D}\,:\,\mathrm{V}_2(x)^p\geq s\}} \alpha_1\big(\mathrm{V}_2(x)^p\big) \leq \inf_{\{x\in\mathcal{D}\,:\,\mathrm{V}_2(x)^p\geq s\}} \mathrm{V}_1(x) = \psi_p(s)\\ &\leq \phi_p(s) = \sup_{\{x\in\mathcal{D}\,:\,\mathrm{V}_2(x)^p\leq s\}} \mathrm{V}_1(x) \leq \sup_{\{x\in\mathcal{D}\,:\,\mathrm{V}_2(x)^p\leq s\}} \alpha_2\big(\mathrm{V}_2(x)^p\big) \leq \alpha_2(s).\end{aligned} \tag{4.12}$$

Now, since α_1,$\alpha_2 \in \mathcal{K}_\infty$ are convex and concave, respectively, it follows by Lemma A.1 that $s^2/\alpha_1(s)$ and $\alpha_2(s)$ are in $\mathcal{O}(s)$ as $s \to \infty$. Knowing from (4.12) that $\alpha_1(s) \leq$

$\psi_p(s) \le \phi_p(s) \le \alpha_2(s)$ for all $s \ge 0$, we conclude that the functions $s^2/\psi_p(s)$ and $\phi_p(s)$ are also in $\mathcal{O}(s)$ as $s \to \infty$, which is the property P4. ∎

The following example shows ways in which the conditions of Theorem 4.6 might fail.

EXAMPLE 4.7. (Illustration of Theorem 4.6). Let $\mathrm{V}_2 : \mathbb{R}^2 \to \mathbb{R}_{\ge 0}$ be the distance to the set $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 = 0\}$, i.e., $\mathrm{V}_2(x_1, x_2) = |x_1|$. Consider the following cases:

P2 fails (ψ_1 is not positive definite): Let $\mathrm{V}_1(x_1, x_2) = |x_1|e^{-|x_2|}$ for $(x_1, x_2) \in \mathbb{R}^2$. Note that V_1 is not $\mathcal{K}$-dominated by V_2 because, given any $\alpha_1 \in \mathcal{K}$, for every $x_1 \in \mathbb{R}$ with $|x_1| > 0$ there exists $x_2 \in \mathbb{R}$ such that the inequality $\alpha_1(|x_1|) \le |x_1|e^{-|x_2|}$ does not hold (just choose x_2 satisfying $|x_2| > \log\big(\frac{|x_1|}{\alpha_1(|x_1|)}\big)$). Thus, there must be some of the hypotheses on Theorem 4.6 that fail to be true. In this case, we observe that

$$\psi_1(s) = \inf_{\{(x_1,x_2)\in\mathbb{R}^2 : |x_1| \ge s\}} |x_1|e^{-|x_2|}$$

is identically 0 for all $s \ge 0$, so it is not positive definite as required in P2.

P2 fails (ϕ_1 is not locally bounded): Let $\mathrm{V}_1(x_1, x_2) = |x_1|e^{|x_2|}$ for $(x_1, x_2) \in \mathbb{R}^2$. As above, one can show that α_2 does not exist in the required class; in this case, the hypothesis P2 is not satisfied because ϕ_1 is not locally bounded in $(0, \infty)$:

$$\phi_1(s) = \sup_{\{(x_1,x_2)\in\mathbb{R}^2 : |x_1| \le s\}} |x_1|e^{|x_2|} = \infty, \quad \forall\, s > 0.$$

P2 fails (ϕ_1 is not right continuous): Let $\mathrm{V}_1(x_1, x_2) = |x_1|^4 + |\sin(x_1x_2)|$ for $(x_1, x_2) \in \mathbb{R}^2$. For every $p > 0$, we have that

$$\phi_p(s) = \sup_{\{(x_1,x_2)\in\mathbb{R}^2 : |x_1|^p \le s\}} |x_1|^4 + |\sin(x_1x_2)| \le s^{4/p} + 1,$$

so ϕ_p is locally bounded in $\mathbb{R}_{\ge 0}$, and, again for every $p > 0$,

$$\psi_p(s) = \inf_{\{(x_1,x_2)\in\mathbb{R}^2 : |x_1|^p \ge s\}} |x_1|^4 + |\sin(x_1x_2)| \ge s^{4/p},$$

so ψ_p is positive definite. However, ϕ_p is not right continuous at 0 because $\sin(x_1x_2) = 0$ when $x_1 = 0$, but $\sup_{\{(x_1,x_2)\in\mathbb{R}^2 : |x_1|^p \le s_0\}} \sin(x_1x_2) = 1$ for any $s_0 > 0$, so by Theorem 4.6 *(i)*, it follows that V_1 is not $\mathcal{K}$-dominated by V_2.

P4 fails (non-compliant asymptotic behavior): Let $\mathrm{V}_1(x_1, x_2) = |x_1|^4$ for $(x_1, x_2) \in \mathbb{R}^2$. Then P2 is satisfied and P3 also holds because $\lim_{s\to\infty} \psi_1(s) = \lim_{s\to\infty} s^4 = \infty$, so Theorem 4.6 *(ii)* implies that V_1 and V_2 are $\mathcal{K}_\infty$-proper with respect to each other. However, in this case $\phi_p(s) = \psi_p(s) = s^{4/p}$, which implies that ϕ_p is not in $\mathcal{O}(s)$ as $s \to \infty$ when $p \in (0, 4)$, and $s^2/\psi_p(s)$ is not in $\mathcal{O}(s)$ as $s \to \infty$ when $p > 4$. Thus P4 is satisfied only for $p = 4$, so Theorem 4.6 *(iii)* implies that only in this case V_1 and V_2^p are $\mathcal{K}_\infty^{cc}$- proper with respect to each other. Namely, for $p > 4$, one cannot choose a convex $\alpha_1 \in \mathcal{K}_\infty$ such that $\alpha_1(|x_1|^p) \le |x_1|^4$ for all $x_1 \in \mathbb{R}$ and, if $p < 4$, one cannot choose a concave $\alpha_2 \in \mathcal{K}_\infty$ such that $|x_1|^4 \le \alpha_2(|x_1|^p)$ for all $x_1 \in \mathbb{R}$. •

4.3. Application to noise-to-state stability. In this section we use the results of Sections 4.1 and 4.2 to study the noise-to-state stability properties of stochastic differential equations of the form (2.1). Our first result provides a way to check whether a candidate function that satisfies a dissipation inequality of the type (3.3)

is in fact a noise-dissipative Lyapunov function, a strong NSS-Lyapunov function in probability, or a pth moment NSS-Lyapunov function.

COROLLARY 4.8. (Establishing proper relations between pairs of functions through seminorms). *Consider* $\mathrm{V}_1, \mathrm{V}_2 : \mathcal{D} \to \mathbb{R}_{\geq 0}$ *such that their nullset is a subspace* $\mathcal{U}$*. Let* $A, \tilde{A} \in \mathbb{R}^{m\times n}$ *be such that* $\mathcal{N}(A) = \mathcal{U} = \mathcal{N}(\tilde{A})$*. Assume that* V_1 *and* V_2 *satisfy* $\{\mathrm{P}i\}_{i=0}^{3}$ *with respect to* $\|.\|_A$ *and* $\|.\|_{\tilde{A}}$*, respectively. Then, for any* $q > 0$*,*

$$\mathrm{V}_1 \sim^{\mathcal{K}_\infty} \mathrm{V}_2, \quad \mathrm{V}_1 \sim^{\mathcal{K}_\infty} \|.\|_A^q, \quad \mathrm{V}_2 \sim^{\mathcal{K}_\infty} \|.\|_{\tilde{A}}^q \quad \textit{in} \quad \mathcal{D}.$$

If, in addition, V_1 *and* V_2 *satisfy* P4 *with respect to* $\|.\|_A$ *and* $\|.\|_{\tilde{A}}$*, respectively, for some* $p > 0$*, then*

$$\mathrm{V}_1 \sim^{\mathcal{K}_\infty^{cc}} \mathrm{V}_2, \quad \mathrm{V}_1 \sim^{\mathcal{K}_\infty^{cc}} \|.\|_A^p, \quad \mathrm{V}_2 \sim^{\mathcal{K}_\infty^{cc}} \|.\|_{\tilde{A}}^p \quad \textit{in} \quad \mathcal{D}.$$

Proof. The statements follow from the characterizations in Theorem 4.6 *(ii)* and *(iii)*, and from the fact that the relations $\sim^{\mathcal{K}_\infty}$ and $\sim^{\mathcal{K}_\infty^{cc}}$ are equivalence relations as shown in Lemma 4.4. That is, under the hypothesis P0,

$$\left.\begin{array}{l}\mathrm{V}_1 \text{ satisfies } \{\mathrm{Pi}\}_{i=1}^{3} \text{ w/ respect to } \|.\|_A \ (\Leftrightarrow\ \mathrm{V}_1 \sim^{\mathcal{K}_\infty} \|.\|_A \text{ in } \mathcal{D})\\ \mathrm{V}_2 \text{ satisfies } \{\mathrm{Pi}\}_{i=1}^{3} \text{ w/ respect to } \|.\|_{\tilde{A}} \ (\Leftrightarrow\ \mathrm{V}_2 \sim^{\mathcal{K}_\infty} \|.\|_{\tilde{A}} \text{ in } \mathcal{D})\end{array}\right\} \Rightarrow \mathrm{V}_1 \sim^{\mathcal{K}_\infty} \mathrm{V}_2 \text{ in } \mathcal{D},$$

$$\left.\begin{array}{l}\mathrm{V}_1 \text{ satisfies } \{\mathrm{Pi}\}_{i=1}^{4} \text{ w/ respect to } \|.\|_A \ (\Leftrightarrow\ \mathrm{V}_1 \sim^{\mathcal{K}_\infty^{cc}} \|.\|_A^p \text{ in } \mathcal{D})\\ \mathrm{V}_2 \text{ satisfies } \{\mathrm{Pi}\}_{i=1}^{4} \text{ w/ respect to } \|.\|_{\tilde{A}} \ (\Leftrightarrow\ \mathrm{V}_2 \sim^{\mathcal{K}_\infty^{cc}} \|.\|_{\tilde{A}}^p \text{ in } \mathcal{D})\end{array}\right\} \Rightarrow \mathrm{V}_1 \sim^{\mathcal{K}_\infty^{cc}} \mathrm{V}_2 \text{ in } \mathcal{D}.$$

Note that, by Lemma 4.3 and (4.1), the equivalences

$$\|.\|_A \sim^{\mathcal{K}_\infty} \|.\|_{\tilde{A}}^q \quad \text{in } \ \mathcal{D}, \qquad \|.\|_A^p \sim^{\mathcal{K}_\infty^{cc}} \|.\|_{\tilde{A}}^p \quad \text{in } \ \mathcal{D}$$

hold for any $p, q > 0$ and any matrices A, $\tilde{A} \in \mathbb{R}^{m\times n}$ with $\mathcal{N}(A) = \mathcal{N}(\tilde{A})$. ∎

We next build on this result to provide an alternative formulation of Corollary 3.9. To do so, we employ the observation made in Remark 4.2 about the possibility of interpreting the candidate functions as defined on a constrained domain of an extended Euclidean space.

COROLLARY 4.9. (The existence of a $p$thNSS-Lyapunov function implies pth moment NSS –revisited). *Under Assumption 2.1, let* $\mathrm{V} \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R}_{\geq 0})$*,* $\mathrm{W} \in \mathcal{C}(\mathbb{R}^n; \mathbb{R}_{\geq 0})$ *and* $\sigma \in \mathcal{K}$ *be such that the dissipation inequality* (3.4) *holds. Let* $R : \mathbb{R}^n \to \mathbb{R}^{(m-n)}$*, with* $m \geq n$*,* $\mathcal{D} \subset \mathbb{R}^m$*,* $\hat{\mathrm{V}} \in \mathcal{C}^2(\mathcal{D}; \mathbb{R}_{\geq 0})$ *and* $\hat{\mathrm{W}} \in \mathcal{C}(\mathcal{D}; \mathbb{R}_{\geq 0})$ *be such that, for* $i(x) = [x^T, R(x)^T]^T$*, one has*

$$\mathcal{D} = i(\mathbb{R}^n), \quad \mathrm{V} = \hat{\mathrm{V}} \circ i, \quad \textit{and} \quad \mathrm{W} = \hat{\mathrm{W}} \circ i.$$

Let $A = \mathrm{diag}(A_1, A_2)$ *and* $\tilde{A} = \mathrm{diag}(\tilde{A}_1, \tilde{A}_2)$ *be block-diagonal matrices, with* $A_1, \tilde{A}_1 \in \mathbb{R}^{n\times n}$ *and* $A_2, \tilde{A}_2 \in \mathbb{R}^{(m-n)\times(m-n)}$*, such that* $\mathcal{N}(A) = \mathcal{N}(\tilde{A})$ *and*

$$\|R(x)\|_{A_2}^2 \leq \kappa \|x\|_{A_1}^2 \tag{4.13}$$

for some $\kappa > 0$*, for all* $x \in \mathbb{R}^n$*. Assume that* $\hat{\mathrm{V}}$ *and* $\hat{\mathrm{W}}$ *satisfy the properties* $\{\mathrm{P}i\}_{i=0}^{4}$ *with respect to* $\|.\|_A$ *and* $\|.\|_{\tilde{A}}$*, respectively, for some* $p > 0$*. Then the system* (2.1) *is NSS in probability and in pth moment with respect to* $\mathcal{N}(A_1)$*.*

Proof. By Corollary 4.8, we have that

$$\hat{\mathrm{V}} \sim^{\mathcal{K}_\infty^{cc}} \hat{\mathrm{W}}, \quad \text{and} \quad \hat{\mathrm{V}} \sim^{\mathcal{K}_\infty^{cc}} \|.\|^p_{\mathrm{diag}(A_1,\, A_2)} \quad \text{in} \quad \mathcal{D}. \tag{4.14}$$

As explained in Remark 4.2, the first relation implies that $\mathrm{V} \sim^{\mathcal{K}_\infty^{cc}} \mathrm{W}$ in $\mathbb{R}^n$. This, together with the fact that (3.4) holds, implies that V is a noise-dissipative Lyapunov function for (2.1). Also, setting $\hat{x} = i(x)$ and using (4.13), we obtain that

$$\|x\|^2_{A_1} \le \|\hat{x}\|^2_{\mathrm{diag}(A_1,\, A_2)} = \|x\|^2_{A_1} + \|R(x)\|^2_{A_2} \le (1+\kappa)\|x\|^2_{A_1},$$

so, in particular, $\|[\,.,R(.)]\|^p_{\mathrm{diag}(A_1,\, A_2)} \sim \|.\|^p_{A_1}$ in $\mathbb{R}^n$. Now, from the second relation in (4.14), by Remark 4.2, it follows that $\hat{\mathrm{V}} \circ i \sim^{\mathcal{K}_\infty^{cc}} \|[\,.,R(.)]\|^p_{\mathrm{diag}(A_1,\, A_2)}$ in $\mathbb{R}^n$. Thus, using (4.1) and Lemma 4.4, we conclude that $\mathrm{V} \sim^{\mathcal{K}_\infty^{cc}} \|.\|^p_{A_1}$ in $\mathbb{R}^n$. In addition, the Euclidean distance to the set $\mathcal{N}(A_1)$ is equivalent to $\|.\|_{A_1}$, i.e., $|.|_{\mathcal{N}(A_1)} \sim \|.\|_{A_1}$. This can be justified as follows: choose $B \in \mathbb{R}^{n\times k}$, with $k = \dim(\mathcal{N}(A_1))$, such that the columns of B form an orthonormal basis of $\mathcal{N}(A_1)$. Then,

$$|x|_{\mathcal{N}(A_1)} = \|(\mathrm{I} - BB^T)x\|_2 = \|x\|_{\mathrm{I}\,-\,BB^T} \sim \|.\|_{A_1}, \tag{4.15}$$

where the last relation follows from Lemma 4.3 because $\mathcal{N}(\mathrm{I} - BB^T) = \mathcal{N}(A_1)$. Summarizing, $\mathrm{V} \sim^{\mathcal{K}_\infty^{cc}} \|.\|^p_{A_1}$ and $\|.\|^p_{A_1} \sim |x|^p_{\mathcal{N}(A_1)}$ in $\mathbb{R}^n$ (because the pth power is irrelevant for the relation $\sim$). As a consequence,

$$\mathrm{V} \sim^{\mathcal{K}_\infty^{cc}} |.|^p_{\mathcal{N}(A_1)} \quad \text{in} \quad \mathbb{R}^n, \tag{4.16}$$

which implies condition (3.21) with convex $\alpha_1 \in \mathcal{K}_\infty$, concave $\alpha_2 \in \mathcal{K}_\infty$, and $\mathcal{U} = \mathcal{N}(A_1)$. Therefore, V is a pth moment NSS-Lyapunov function with respect to the set $\mathcal{N}(A_1)$, and the result follows from Corollary 3.9. ∎

5. Conclusions. We have studied the stability properties of SDEs subject to persistent noise (including the case of additive noise). We have generalized the concept of noise-dissipative Lyapunov function and introduced the concepts of strong NSS-Lyapunov function in probability and pth moment NSS-Lyapunov function, both with respect to a closed set. We have shown that noise-dissipative Lyapunov functions have NSS dynamics and established that the existence of an NSS-Lyapunov function, of either type, with respect to a closed set, implies the corresponding NSS property of the system with respect to the set. In particular, pth moment NSS with respect to a set provides a bound, at each time, for the pth power of the distance from the state to the set, and this bound is the sum of an increasing function of the size of the noise covariance and a decaying effect of the initial conditions. This bound can be achieved regardless of the possibility that inside the set some combination of the states accumulates the variance of the noise. This is a meaningful stability property for the aforementioned class of systems because the presence of persistent noise makes it impossible to establish in general a stochastic notion of asymptotic stability for the set of equilibria of the underlying differential equation. We have also studied in depth the inequalities between pairs of functions that appear in the various notions of Lyapunov functions mentioned above. We have shown that these inequalities define equivalence relations and have developed a complete characterization of the properties that two functions must satisfy to be related by them. Finally, building on this characterization, we have provided an alternative statement of our stochastic stability results. Future work will include the study of the effect of delays and impulsive right-hand sides in the class of SDEs considered in this paper.

Acknowledgments. The first author would like to thank Dean Richert for useful discussions. In addition, the authors would like to thank Dr. Fengzhong Li for his kind observations that have made possible an important correction of the proof of Theorem 3.6. The research was supported by NSF award CMMI-1300272.

References.

[1] V. S. Borkar, *Probability Theory: An Advanced Course*, Springer, New York, 1995.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2009.

[3] H. Deng and M. Krstić, *Output-feedback stabilization of stochastic nonlinear systems driven by noise of unknown covariance*, Systems & Control Letters, 39 (2000), pp. 173–182.

[4] H. Deng, M. Krstić, and R. J. Williams, *Stabilization of stochastic nonlinear systems driven by noise of unknown covariance*, IEEE Transactions on Automatic Control, 46 (2001), pp. 1237–1253.

[5] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[6] H. K. Khalil, *Nonlinear Systems*, Prentice Hall, 3 ed., 2002.

[7] R. Khasminskii, *Stochastic Stability of Differential Equations*, vol. 66 of Stochastic Modelling and Applied Probability, Springer, 2012.

[8] F. C. Klebaner, *Introduction to Stochastic Calculus With Applications*, Imperial College Press, 2005.

[9] P. E. Kloeden and T. Lorenz, *Mean-square random dynamical systems*, Journal of Differential Equations, 253 (2012), pp. 1422–1438.

[10] V. Lakshmikantham, V. M. Matrosov, and S. Sivasundaram, *Vector Lyapunov Functions and Stability Analysis of Nonlinear Systems*, vol. 63 of Mathematics and its Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.

[11] J. Lewin and M. Lewin, *An Introduction to Mathematical Analysis*, BH mathematics series, The Random House, 1988.

[12] S.-J. Liu, J.-F. Zhang, and Z.-P. Jiang, *A notion of stochastic input-to-state stability and its application to stability of cascaded stochastic nonlinear systems*, Acta Mathematicae Applicatae Sinica, English Series, 24 (2008), pp. 141–156.

[13] X. Mao, *Stochastic versions of the LaSalle theorem*, Journal of Differential Equations, (1999), pp. 175–195.

[14] ———, *Stochastic Differential Equations and Applications*, Woodhead Publishing, 2nd ed., 2011.

[15] D. Mateos-Núñez and J. Cortés, *Stability of stochastic differential equations with additive persistent noise*, in American Control Conference, Washington, D.C., June 2013, pp. 5447–5452.

[16] J. N. McDonald and N. A. Weiss, *A Course in Real Analysis*, Elsevier, Oxford, UK, 1999.

[17] J. R. Movellan, *Tutorial on stochastic differential equations*, tutorial, MPLab, UCSD, 2011.

[18] B. Øksendal, *Stochastic Differential Equations - An Introduction with Applications*, Universitext, Springer-Verlag, 2010.

[19] L. Praly and Y. Wang, *Stabilization in spite of matched unmodeled dynamics and an equivalent definition of input-to-state stability*, Mathematics of Control,

Signals and Systems, 9 (1996), pp. 1–33.
[20] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
[21] H. Schurz, *On moment-dissipative stochastic dynamical systems*, Dynamic systems and applications, 10 (2001), pp. 11–44.
[22] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, vol. 6 of TAM, Springer, 2 ed., 1998.
[23] ———, *Input to state stability: Basic concepts and results*, Nonlinear and Optimal Control Theory, 1932 (2008), pp. 163–220.
[24] E. D. Sontag and Y. Wang, *On characterizations of the input-to-state stability property*, Systems & Control Letters, 24 (1995), pp. 351–359.
[25] ———, *New characterizations of input-to-state stability*, IEEE Transactions on Automatic Control, 41 (1996), pp. 1283 – 1294.
[26] T. Taniguchi, *The asymptotic behaviour of solutions of stochastic functional differential equations with finite delays by the Lyapunov-Razumikhin method*, in Advances in Stability Theory at the End of the 20th Century, A. A. Martynyuk, ed., vol. 13 of Stability and Control: Theory, Methods and Applications, Taylor and Francis, 2003, pp. 175–188.
[27] U. H. Thygesen, *A survey of Lyapunov techniques for stochastic differential equations*, Tech. Rep. 18-1997, Department of Mathematical Modeling, Technical University of Denmark, 1997.
[28] F. Wu and P. E. Kloeden, *Mean-square random attractors of stochastic delay differential equations with random delay*, Discrete and Continuous Dynamical Systems - Series B (DCDS-B), 18 (2013), pp. 1715–1734.
[29] Z.-J. Wu, X.-J. Xie, and S.-Y. Zhang, *Adaptive backstepping controller design using stochastic small-gain theorem*, Automatica, 43 (2007), pp. 608–620.

Appendix. The next result is used in the proof of Theorem 4.6.

Lemma A.1. (Existence of bounding convex and concave functions in $\mathcal{K}_\infty$). *Let α be a class $\mathcal{K}_\infty$ function. Then the following are equivalent:*

(i) There exist $s_0 \geq 0$ and $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$, convex and concave, respectively, such that $\alpha_1(s) \leq \alpha(s) \leq \alpha_2(s)$ for all $s \geq s_0$, and
(ii) $\alpha(s)$, $s^2/\alpha(s)$ are in $\mathcal{O}(s)$ as $s \to \infty$.

Proof. The implication *(i)* $\Rightarrow$ *(ii)* follows because, for any $s \geq s_0 > 0$,

$$\tfrac{\alpha_1(s_0)}{s_0} s \leq \alpha_1(s) \leq \alpha(s) \leq \alpha_2(s) \leq \tfrac{\alpha_2(s_0)}{s_0} s,$$

by convexity and concavity, respectively, where $\alpha_1(s_0), \alpha_2(s_0) > 0$.

To show *(ii)* $\Rightarrow$ *(i)*, we proceed to construct α_1, α_2 as in the statement using the correspondence between functions, graphs and epigraphs (or hypographs). Let $\alpha_1 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be the function whose epigraph is the convex hull of the epigraph of α, i.e., $\operatorname{epi} \alpha_1 \triangleq \operatorname{conv}(\operatorname{epi} \alpha)$. Thus, α_1 is convex, nondecreasing, and $0 \leq \alpha_1(s) \leq \alpha(s)$ for all $s \geq 0$ because $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \supseteq \operatorname{epi} \alpha_1 = \operatorname{conv}(\operatorname{epi} \alpha) \supseteq \operatorname{epi} \alpha$. Moreover, α_1 is continuous in $(0, \infty)$ by convexity [20, Th. 10.4], and is also continuous at 0 by the sandwich theorem [11, p. 107] because $\alpha \in \mathcal{K}_\infty$. To show that $\alpha_1 \in \mathcal{K}_\infty$, we have to check that it is unbounded, positive definite in $\mathbb{R}_{\geq 0}$, and strictly increasing. First, since $s^2/\alpha(s) \in \mathcal{O}(s)$ as $s \to \infty$, there exist constants $c_1, s_0 > 0$ such that $\alpha(s) \geq c_1 s$ for all $s > s_0$. Now, define $g_1(s) \triangleq \alpha(s)$ if $s \leq s_0$ and $g_1(s) \triangleq c_1 s$ if $s > s_0$, and $g_2(s) \triangleq -c_1 s_0 + c_1 s$ for all $s \geq 0$, so that $g_2 \leq g_1 \leq \alpha$. Then,

$\mathrm{epi}\,\alpha_1 = \mathrm{conv}(\mathrm{epi}\,\alpha) \subseteq \mathrm{conv}(\mathrm{epi}\,g_1) \subseteq \mathrm{epi}\,g_2$, because $\mathrm{epi}\,g_2$ is convex, and thus α_1 is unbounded. Also, since $\mathrm{conv}(\mathrm{epi}\,g_1) \cap \mathbb{R}_{\geq 0} \times \{0\} = \{(0,0)\}$, it follows that α_1 is positive definite. To show that α_1 is strictly increasing, we use two facts: since α_1 is convex, we know that the set in which α_1 is allowed to be constant must be of the form $[0, b]$ for some $b > 0$; on the other hand, since α_1 is positive definite, it is nonconstant in any neighborhood of 0. As a result, α_1 is nonconstant in any subset of its domain, so it is strictly increasing.

Next, let $\alpha_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be the function whose hypograph is the convex hull of the hypograph of α, i.e., $\mathrm{hyp}\,\alpha_2 \triangleq \mathrm{conv}(\mathrm{hyp}\,\alpha)$. The function α_2 is well-defined because $\alpha(s) \in \mathcal{O}(s)$ as $s \to \infty$, i.e., there exist constants $c_2, s_0 > 0$ such that $\alpha(s) \leq c_2 s$ for all $s > s_0$, so if we define $g(s) \triangleq c_2 s_0 + c_2 s$ for all $s \geq 0$, then $\mathrm{hyp}\,\alpha_2 = \mathrm{conv}(\mathrm{hyp}\,\alpha) \subseteq \mathrm{hyp}\,g$, because $\mathrm{hyp}\,g$ is convex, and thus $\alpha_2(s) \leq g(s)$. Also, by construction, α_2 is concave, nondecreasing, and $\alpha_2 \geq \alpha$ because $\mathrm{hyp}\,\alpha_2 \supseteq \mathrm{hyp}\,\alpha$, which also implies that α_2 is unbounded. Moreover, α_2 is continuous in $(0, \infty)$ by concavity [20, Th. 10.4], and is also continuous at 0 because the possibility of an infinite jump is excluded by the fact that $\alpha_2 \leq g$. To show that $\alpha_2 \in \mathcal{K}_\infty$, we have to check that it is positive definite in $\mathbb{R}_{\geq 0}$ and strictly increasing. Note that α_2 is positive definite because $\alpha_2(0) = 0$ and $\alpha_2 \geq \alpha$. To show that α_2 is strictly increasing, we reason by contradiction. Assume that α_2 is constant in some closed interval of the form $[s_1, s_2]$, for some $s_2 > s_1 \geq 0$. Then, as α_2 is concave, we conclude that it is nonincreasing in (s_2, ∞). Now, since α_2 is continuous, we reach the contradiction that $\lim_{s\to\infty} \alpha(s) \leq \lim_{s\to\infty} \alpha_2(s) \leq \alpha_2(s_1) < \infty$. Hence, α_2 is strictly increasing. $\square$